JAN 24 1924
PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

Journal of Educational Research

OFFICIAL ORGAN OF THE
EDUCATIONAL RESEARCH ASSOCIATION

B. R. BUCKINGHAM, EDITOR

ASSOCIATE EDITORS

E. J. ASHBAUGH — W. W. CHARTERS — GEORGE D. STRAYER
GUY M. WHIPPLE — S. A. COURTIS — LEWIS M. TERMAN
HARRY A. GREENE

VOL. IX JANUARY, 1924 No. 1

CONTENTS

	Page
SIZE OF CLASS AS A FACTOR IN UNIVERSITY INSTRUCTION — J. B. Edmonson and F. J. Mulder	1
WHAT SHOULD CHILDREN READ IN SCHOOL? — B. R. Buckingham	13
THE CHALLENGE OF THE PRIVATE SCHOOL — Norman Fenton and Lowry S. Howard	22
WHAT IS MEASURED BY INTELLIGENCE TESTS? — Omen Bishop	29
THE SPEED FACTOR IN MENTAL MEASUREMENTS — Giles Murrel Ruch	39
A STUDY OF INTELLIGENCE SCALES FOR GRADES TWO AND THREE — J. Cayce Morrison, W. B. Cornell and Ethel Cornell	46
PROPOSALS FOR STANDARDIZING MEASUREMENT IN EDUCATION. II. THE COEFFICIENT OF APPLICATION — Harry S. Will	57
A SERVICEABLE P. E. TABLE — Herbert A. Toops and Zaida F. Miner	63
EDITORIALS	69
REVIEWS AND ABSTRACTS	73
NEWS ITEMS AND COMMUNICATIONS	85
EDUCATIONAL RESEARCH ASSOCIATION	91

Entered as second-class matter at the Post Office at Bloomington, Illinois under the act of March 3, 1879.

Journal of Educational Research

Copyright 1923, by
Public School Publishing Co.

Official Organ of the Educational Research Association

Editor-in-Chief, B. R. Buckingham
Director of Bureau of Educational Research, Ohio State University, Columbus, Ohio

Associate Editors

E. J. ASHBAUGH, *Ohio State University, Columbus.*

W. W. CHARTERS, *Carnegie Inst. Tech., Pittsburgh.*

S. A. COURTIS, *City Schools, Detroit, Mich.*

HARRY A. GREENE, *Ex-Officio, Secretary Educational Research Association, State University of Iowa, Iowa City.*

GEORGE D. STRAYER, *Teachers College, Columbia University, New York City.*

LEWIS M. TERMAN, *Leland Stanford Junior University, Stanford University, California*

GUY M. WHIPPLE, *University of Michigan, Ann Arbor, Michigan.*

Contributing Editors

VIRGIL E. DICKSON, *Director, Bureau of Research and Guidance, Public Schools, Oakland and Berkeley, California.*

CHARLES FORDYCE, *Director of Bureau of Educational Measurements, University of Nebraska, Lincoln, Nebr.*

M. E. HAGGERTY, *Dean of the College of Education, University of Minnesota, Minneapolis.*

V. A. C. HENMON, *Director, School of Education, University of Wisconsin, Madison.*

ARTHUR W. KALLOM, *Assistant Director, Department of Educational Investigation and Measurement, Public Schools, Boston, Mass.*

F. J. KELLY, *Dean of University Administration, University of Kansas, Lawrence, Kans.*

ROBT. H. LANE, *Assistant Superintendent of Schools Los Angeles, Cal.*

WILLIAM A. McCALL, *Assistant Professor of Education, Teachers' College, Columbia University.*

GEORGE W. MELCHER, *Director, Bureau of Research and Efficiency, Public Schools, Kansas City, Missouri.*

EUGENE A. NIFENECKER, *Assistant Director, Bureau of Reference, Research and Statistics, Public Schools, New York City.*

JOSEPH P. O'HERN, *Assistant Superintendent of Schools and Director of Efficiency Bureau, Public Schools, Rochester, New York.*

LEON O. SMITH, *Assistant Superintendent (in charge of Educational Research) Public Schools, Omaha, Nebraska.*

W. W. THEISEN, *Assistant Superintendent of Schools Milwaukee, Wis.*

CLIFFORD WOODY, *Director Bureau of Tests and Measurements, University of Michigan, Ann Arbor, Mich.*

The JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except July and August by the Public School Publishing Co. The subscription price is $4.00 $4.20 in Canada or Mexico; and $4.50 in other foreign countries.

The editorial office is at Columbus, Ohio. Address all communications on editorial matters to B. R. Buckingham, Director of Bureau of Educational Research, The Ohio State University, Columbus, Ohio.

Subscriptions and all communications about reprints or other business matters should be addressed to the Public School Publishing Company at Bloomington, Illinois.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume IX JANUARY, 1924 Number 1

SIZE OF CLASS AS A FACTOR IN UNIVERSITY INSTRUCTION

J. B. EDMONSON and F. J. MULDER
University of Michigan

During the first semester of the school year 1922-23 an experiment was conducted by the authors to determine what differences existed in the measurable results of instruction in a large and a small class in an education course offered in the School of Education of the University of Michigan. The two classes were taught by the same instructor, and the course was entitled *An introduction to high-school problems.* The achievements of students in these classes were in terms of the semester averages of the classes as measured by written reports, short written quizzes, and semester examinations. In order to form a basis for our judgment as to what might constitute a large and a small class, a distribution of the classes in the School of Education for the first semester of 1922-23 was made according to size. The results are shown in Table I.

The range it will be seen was from 19 to 124 in a class with 65 as the median. Four of these classes were in the subject *An introduction to high-school problems* and the range in the sizes of these four classes was from 45 to 109. Since the smallest had an enrolment well below the median of 65 and the largest an enrolment well above the median, these two classes may properly be called large and small classes for the group to which they belong.

Equating of groups.—An attempt was made to arrange such conditions for the experiment that the size of class would be

TABLE I. DISTRIBUTION OF CLASSES IN THE SCHOOL OF EDUCATION

Number in classes	Frequency
15— 29	5
30— 44	2
45— 59	3
60— 74	3
75— 89	5
90—104	2
105—119	1
120—134	1
Total	22
Q_1	27
Median	65
Q_3	88
Range	19–124

the only variable. In order that we might obtain comparable results, the plan was followed of giving intelligence tests to both classes, pairing one member of the larger class with a member of the smaller class, and using only the achievements of the paired groups in the comparison.

The intelligence test used was the scrambled Alpha. In order to make the results as nearly equal as possible, two questions were asked at the time of giving the tests. Each student was asked, first, whether or not he had taken an intelligence test before, and second, if he had taken one before, whether or not he had taken this particular intelligence test. The answers to these questions were recorded after the score made by each student. Then for each member of the smaller class a member of the larger class was found who not only had the same or almost the same intelligence score but whose status as far as having taken intelligence tests before was also the same. The method of pairing in this manner compensated for any increase in score that might occur due to familiarity with intelligence tests in general or with the scrambled Alpha test in particular.

Table II shows the distribution of scores for the large class, the small class, and the selected group from the large class. The

differences between the numbers of students in the classes here and the numbers noted before are accounted for by absences on the day the intelligence tests were given. It will be seen from Table II that the two groups used in this experiment are nearly equal in intelligence; the medians, in fact being exactly the same, 148.7.

TABLE II. DISTRIBUTION OF SCORES FOR THE VARIOUS GROUPS

Scores	Students of Large Class	Students of Small Class	Selected Group from Large Class
90— 94	1		
95— 99	1		
100—104	2		
105—109	2	1	1
110—114	5	0	4
115—119	2	6	1
120—124	2	3	1
125—129	5	2	5
130—134	8	3	3
135—139	9	3	3
140—144	16	2	2
145—149	7	2	2
150—154	9	6	5
155—159	13	6	8
160—164	4	4	2
165—169	7	5	5
170—174	3	0	1
175—179	1		
180—184			
Totals.....	97	43	44
1st Quartile...	132.7	126.8	128.7
Median.....	143.6	148.7	148.7
3rd Quartile...	156.4	158.1	158.2

Admission to these classes was limited to students without teaching experience so that the classes were comparable in this respect. The number of men in the classes was small—12 in the large-class group and 10 in the small class. The proportions were therefore sufficiently similar to prevent any material differences due to sex. Thus we have two groups to all appearances equal, but one taught in a relatively small class and the other taught in a class almost two and a half times as large. It is

important to emphasize that the group of 43 students which we refer to as the large-class group was never taken out of the class of 109 nor was either class apprised of the fact that this investigation was being made until the semester's work was nearly completed.

Content of the course.—The two classes were conducted as quiz and discussion classes and the work assigned was always the same for both classes. This work consisted of assignments in textbooks, references to library books and magazines, and short papers or essays covering major topics in the course. Occasionally a period was occupied by the instructor with a lecture or discussion of various new developments in the field of education. Two textbooks were used—Colvin's *An Introduction to High-School Teaching* and Lewis' *Democracy's High School.* The report of the N.E.A. Commission on the Reorganization of Secondary Education entitled *Cardinal Principles of Secondary Education* was also used. Several times during the semester men from other departments were called in to lecture on some topic directly connected with their field of specialization. In this way the course was made a broad, general introduction to the subject of high-school teaching.

Technic of instruction.—The procedure in conducting the two classes is worthy of considerable attention. One of the problems in large groups conducted as discussion classes, is of course to get a reaction from each member of the class. The method often employed by the instructor in the larger section was to state a specific question which was to be discussed and then to have each member of the class note on paper the general nature of the response he would make, or perhaps list his solutions to the question. These papers were seldom collected but the discussion of this question was conducted on the basis of the answers the students had noted on the paper. These were frequently referred to as "reaction questions."

It was at first intended that the procedure in the two sections should be in every respect the same. There was the same in-

structor and the same assigned work in both classes. All papers were marked by the same assistant. Both classes met in the afternoon—the larger at one o'clock and the smaller at two o'clock. The larger class, however, met in a large lecture room capable of seating several hundred people while the smaller class occupied a room which would seat but a few more than the number in the class. The greater compactness and physical nearness of this latter group undoubtedly had an influence on their attitude toward the class. It also became apparent that the difference in the size of the rooms combined with the difference in the size of the classes had some influence on the procedure of the instructor.

In order to obtain some idea of the influence of these factors in the two classes, volunteers were called for to change from the large to the small class for the last few weeks of the semester in order to note differences in procedure. Five students responded to this call; and their report was accepted in lieu of a part of the final examination in the course. None of these students was in the selected group from the large class which was used in the comparison with the small class. Independently of this the instructor also noted differences in conducting the two classes.

The following is a summary of all the differences noted between the two classes by the five transferred pupils:

	Small Class	Large Class
	Physical Differences	
Enrolment—	45 students.	109 students.
Room—		
Size	Small compact room; seated about 60.	Large lecture room; seated about 300.
Location	Room on ground floor. Some disturbance from people passing by on walk just outside of windows.	Room on second floor; poor acoustic properties making it difficult to hear all recitations. Ventilation and lighting poor; in winter steam radiators contributed to disturbing factors.
Seats—	Seats of table recitation type.	Auditorium seats difficult to get in and out of.
Time—	Two o'clock.	One o'clock.

Small Class	Large Class
Differences in Method	
The class was conducted in a less formal manner.	More formal methods.
There was more class discussion.	Little discussion.
There were fewer short written quizzes. Students were not required to list answers to questions as often. (Referred to above as reaction questions.)	More drill.
Individual questions used more often than general class questions.	
Practically all members of class were called on to recite.	More lectures.
Instructor introduced more illustrative material.	
Instructor was less aggressive. Usually remained seated.	Impersonal attitude.
Instructor was better acquainted and took more of a personal interest in each student.	
Differences in Class Response	
Students more mentally alert.	Students mentally sluggish.
No lagging of interest toward end of period.	Only small portion took active part in discussions.
More questions and problems brought up by students.	Majority too frightened to offer remarks unless called upon.
Attitude more studious and attentive.	Took attitude of "Let George do it."
More original comments showing independent thinking.	
More interest and less confusion than in large section.	
The men entered into the discussion with more freedom.	

One other difference which should be noted is the fact that in the larger section an assistant was present who took the roll each day and assisted in passing out and collecting papers. He also read and marked all papers from both sections. He was frequently consulted in regard to work missed during absences, standing in class, assignments, and references. The assistant was seldom present at the meetings of the smaller class.

While each one of these differences by itself might seem trivial and inconsequential, taken all together they seem to pre-

sent rather striking evidence of the fact that the instructor adapted himself and his teaching to the size of the class. Every difference seems to follow as a direct result of the conditions with which he was dealing.

A criticism frequently directed against comparing the achievements of large and small classes is that the same methods of instruction are used in both classes and that therefore one of them may be handicapped by methods poorly adapted to that size of class. The ideal experiment would be to compare results in classes where the method had been adapted to the size of class; that is, where the small class had been taught by small-class methods and the large class by large-class methods. We are inclined to believe that the proper interpretation of the differences in method noted above indicates that we have approached this ideal experiment. We are frank to admit that a person who was not present in the classes might have drawn the opposite conclusion, namely that the failure to arouse in the large class as great an interest as in the small class was due to lack of the use of proper methods. Actual observation in the two classes and a critical study of the differences in method seem to us to point to the first conclusion. On the other hand, one of the values of the study for us is that it points the way to a refinement of methods which may go far to eliminating the objections to the large class.

Marking.—The basis on which the achievement of the classes was measured consisted of the marks on the occasional essays mentioned above, several short objective quizzes, and a mid-semester and final examination. An effort was put forth to make all tests as objective as possible. The mid-semester examination was of the completion type, while the final examination was divided into three parts, the first consisting of "yes and no" questions, the second of completion questions, and the third of a few of the essay type of examination questions.

As we have said before, the marking of all papers for both classes was done by the same assistant. On items in which the judgment of the reader played a part, one or two questions at

a time were marked in all examination papers before the next question was taken up. This was done in order that differences in subjective attitude might affect the results as little as possible.

All of the short essays of the class were marked in terms of the letter ratings, A, B, C, D, E. Later for the purposes of this study a definite numerical value was assigned to each letter.

For all other work—short quizzes, mid-semester examinations, and final examinations—the papers were marked in percents. For determining the semester mark it was felt that the mid-semester and final examinations should be given more weight than the short quizzes and essays. Accordingly, the mid-semester ratings were multiplied by two and the final ratings by three, while all the other marks were multiplied by one. All of the marks for each student were then averaged to obtain his semester average.

The semester marks for the students of each of the paired groups were then averaged to obtain the semester average for the group taught in the small class and the group taught in the large class. A comparison of these averages for the two groups appears in Table III. The averages for each section for the various kinds of work were also determined to see wherein any differences might occur. This summary appears in Table IV.

What differences do exist here indicate that the small class was slightly superior in the essay work and the mid-semester examination while the large class group excelled in the quizzes and final examination.

A comparison of the achievements of the two classes.—Table III shows the comparative achievements of the two groups. It will be seen that what slight differences are revealed—0.41 of a point in the average and 1.4 points in the median—are in favor of the group taught in the large class. The differences in our opinion are too slight to be in any way significant. There is approximately the same range and the same distribution of marks. To all appearances the students in one class obtained, as far as achievement can be measured, just as much as those in the other class.

TABLE III. COMPARISON OF SEMESTER AVERAGES OF THE TWO CLASSES

	Semester Mark		First Quartile	Third Quartile	Range
	Average	Median			
Large Class.....	83.98	84.8	80.5	88	70—93.9
Small Class.....	83.57	83.4	80.2	87.7	72.2—92.7

TABLE IV. COMPARISON OF AVERAGE MARKS FOR DIFFERENT KINDS OF WORK FOR THE TWO CLASSES

	Essay Work		Quizzes		Mid-Semester		Final		Semester Average	
	Average	Range	Average	Range	Average	Range	Average	Range	Average	Range
Large Class	87.3	70–100	77.5	28–100	86.7	68–98	83.4	66–93	83.98	70–93.9
Small Class	90.1	70–100	75.4	20–100	89.1	56–98	79.2	53–96	83.57	72–92.7

Just how much reliance can be placed on the comparison of these two averages is doubtful. There can be no question but that the subjective element entered to a slight extent into the marking of the examination papers and to a larger extent in the marking of the essays. An attempt was made, of course, to reduce this element to a minimum by using the objective quizzes and by the objective character of the mid-semester and final examinations. It could not of course be eliminated from the essay marking. On the other hand, since all of the marking for both classes was done by the same person there would be a tendency for the differences due to this element to disappear in the averages.

Teaching load as a factor.—There are other factors which will of necessity enter into a discussion of class-size.

Were we to conclude that the achievement would be the same whether a person were taught in a small or a large class, we should still have to consider the effect on the teaching load. Teaching work in the university is assigned on the basis of semester hours of teaching regardless of the number of students

taught during any of these hours. It is quite obvious, therefore, that if an instructor does do all the reading of papers and marking of examinations, his teaching load is almost in direct proportion to the number of students in his classes. On the other hand, if this reading work is done to a large extent by assistants, as was true in this case, the work for the instructor is very little more for a large class than for a small class. In fact a large class may very conceivably be easier to teach with the help of an assistant than a small class without such help. The question of size of class, then, will enter into the matter of teaching load to the extent of deciding whether or not the instructor's load is sufficiently great to require an assistant.

Student opinion.—A second consideration to be taken account of is the attitude of the student himself. This element does not enter into the question for elementary schools; but when we come to the much more mature university student who does not attend under compulsion but voluntarily and for a direct personal advantage, it may very well be an important item to consider. In this instance the student attitude was considered important for two reasons: first, because it is recognized that only a part of the total educational product is measured by the means used to obtain the final averages in this experiment; and second, because there are apparently some very definite things which a student wishes to get from a course and if the course does not give him those things it has to a certain extent failed.

An attempt was therefore made during the first part of the semester to obtain some indication of the student opinion on the efficiency of large and small classes in general by a vote in the two classes. The students were asked to think of two, three, or more classes in which they had been previously enrolled and which might possibly be regarded as small or large, that is forty to forty-five or less for a small class and ninety to one hundred or more for a large class. Then they were to rate these classes as either efficient, fairly efficient, or inefficient. The tabulated results are shown in Table V.

TABLE V. RELATIVE EFFICIENCY OF LARGE AND SMALL CLASSES ACCORDING TO THE VOTE OF STUDENTS

	Number of Classes	Efficient		Fair		Inefficient	
		Number	Percent	Number	Percent	Number	Percent
Large classes....	277	132	48	73	26	72	26
Small classes....	272	150	55	65	25	57	20

According to Table V, the students favor the small class on the basis of efficiency but not by a marked margin. The summary classifies as inefficient 20 percent of the small classes as against 26 percent of the large classes. It represents 55 percent of the small classes as efficient and 48 percent of the large classes. Obviously a great many variable factors besides the judgment of the students might have influenced the efficiency of these classes.

As a further check on student opinion concerning size of class, the preferences of all the members of the large class in the previous experiment were secured through a questionnaire. Not having attended a small class under the same instructor they were, of course, not able to base their judgment on actual comparison. The preferences were as follows: Number favoring small classes, 50; number favoring large classes, 14; number believing there was no difference, 34. The reasons given for favoring the small classes together with the frequency of mention were as follows:

More personal relationship with instructor	28
More chance to make contributions in class	12
Instructor can judge marks better	7
Makes one nervous to recite before large class	6
Greater concentration possible in small class	3
Greater interest and attention in small group	5
Large class wastes time	2
Unable to hear all that is said in large room[1]	18
A few individuals in a large section monopolize the discussion	5

[1] This was undoubtedly due to the poor appointment of the particular room in which this class was held.

While a plurality is thus shown in favor of the small class and comparatively few actually favored the large class, there is by no means complete unanimity. The five students who attended both classes, however, and who thus had a basis for direct comparison were unanimously in favor of the small class.

Conclusions.—First, results in terms of semester averages indicate no appreciable difference in the achievement of the large and small class included in this study. Second, student opinion as noted in the questionnaires on the relative efficiency of large and small classes, and on the individual preferences of the students, favors, in both cases, the small class. Third, the chief advantages in a small class appear to be more personal contact with the instructor, more chance to participate in the class discussion and introduction of more varied illustrative material. Fourth, the chief advantages of a large class from the student point of view are the varied opinions gained by contact with so large a group. Fifth, the losses encountered in the large class in this experiment due to poor appointment of the classroom and inability to hear can undoubtedly be greatly reduced. Sixth, there is reason to believe that methods can be devised to overcome the lack of acquaintance between instructor and student in the large classes, which is the chief objection of the students. Seventh, there is need for more study and experimentation in improving the large-class methods in such things as:

1. Methods of obtaining response from each member of the group.
2. Methods of securing personal contact between instructor and class.
3. Methods of maintaining the interest of the entire group in the discussions.

Two unsolved problems suggested by this study are: (1) Do the semester averages based on written work measure the major part of the desired educational product of university instruction? (2) Given an assistant to take care of the paper work and routine matters, is the real teaching load of the instructor greatly affected by increasing the size of class in a discussion or lecture course from forty students to approximately one hundred and twenty?

WHAT SHOULD CHILDREN READ IN SCHOOL?

B. R. Buckingham
Ohio State University

Up to about fifty years ago the prevalent practice in teaching reading was to begin with the alphabet. There was logic in this, for were not thoughts expressed in sentences and these composed of words and phrases; and were not words and phrases, in their last analysis, made up of letters? The most elemental symbols were therefore the letters of the alphabet. I presume that the alphabet method of teaching reading affords the best example of the extent to which logical analysis plays havoc with the teaching process. In the days of our fathers and grandfathers the child on going to school began with *a*. Now, can anyone find anything more abstract than *a*—anything more distressingly remote from the experience of a little child? Yet, according to the logic of teaching, not only *a* but the entire alphabet must be learned before any words could be taught. Not only so, but certain meaningless combinations—*a–b*, ab; *i–b*, ib; *u–b*, ub—must also be learned. Thus the child usually spent the entire year learning the alphabet, and he failed even then to make a very good job of it. For example, the minutiae of eye-fixation required in distinguishing *b* from *d* and *p* from *q* introduced a special difficulty. Vestiges of this difficulty have survived in the language. The anxious mother giving a last admonition to the child as he left to go to school would say, "Mind your *p's* and *q's*." This phrase has been embalmed in the language, though its origin is probably seldom remembered.

In the days of the alphabet method the reading material naturally conformed to the method. The child possessed a hornbook. Now, the hornbook was not a book at all, but a flat, thin board, usually covered with paper, on which were printed the alphabet, a few combinations of letters, and usually the digits, together with the Lord's Prayer. The whole was covered with a transparent piece of horn. This was the child's reader for the

first year or even more. I have said before that he spent the year learning the alphabet. It is clear that if he were going to do this, he would not need anything more than his hornbook. Indeed, the richness of material presented by the hornbook was excessive, for to the alphabet it added the digits and the Lord's Prayer. The latter, however, was committed to memory; and although we may fancy the child gazing at the printed lines, we are quite certain that no reading act took place.

As the method of teaching reading developed, a parallel development in content took place. Doctor Dunn in her admirable study of *Interest Factors in Primary Reading Material* has classified reading textbooks into four chronological periods: (1) those between 1824 and 1867, (2) those between 1868 and 1887, (3) those between 1888 and 1905, and (4) those from 1905 to the present time. The books of the earliest of these periods became rivals and later successors to the hornbook. It is curious to note how difficult it was to break away from the analytical notion—the notion which led people to think that because the letters of the alphabet were the simplest language elements, they were therefore the material for beginners. The first sentences of all the primers of these ante-bellum days were made-up words of two or three letters. No matter how unusual the combination turned out to be, the virtue lay in the fact that each word had not more than three letters. Typical sentences are: "My son, do no ill." "Go not in the way of bad men." "Bad men go to the pit." The more ordinary expression for the destination of bad men obviously could not be used because it exceeds the three-letter limit. The following illustrate the authors' ingenuity in making sentences of two-letter words: On we go. So do we. By me, no. So at me. Go to it. Do as we do. Is my ox to go in to it? Oh, no, it is I to go up.

In these readers, however, there was a very rapid recovery from this condition. One of the readers from which I have been quoting introduces on pages 26 and 27, the following gems: "Vain persons are full of the allurements of dress." "The philosophical infidel can never refer to him for authority." One

is tempted to attribute to the teaching of reading from this book a marvelous rate of progress. If in twenty or thirty pages we may pass from two-letter words to vain persons and their allurements of dress to say nothing of philosophical infidels who can never refer to somebody for authority, something unusual must have happened in the interim. Of course, nothing very wonderful happened. The children ground away at these impossible tasks until by dint of sheer longevity, the best of them learned to read.

The more advanced books of these days ran to moral stories, devotional selections, or such pieces as "The Funeral of a Little Boy," "The Dying Christian to His Soul," and "On the Shortness of Human Life."

It is when we contemplate the situation in those days that we realize the progress we have made. This will give me an opportunity to record unmistakably my sense of the splendid achievement which has been made in the teaching of reading. If in what I may say later I offer some criticism of existing conditions, these criticisms are to be interpreted with reference to the advance which teachers and textbook writers have made—an advance which I am glad to point out. The fact is that the type of criticism to which modern reading materials may be subjected could not have been applied to the materials in use fifty or sixty years ago.

In brief, the period from 1868 to 1887 was characterized so far as reading books were concerned by made-up sentences, and by moralizing selections. Literary materials just began to emerge. They became decidedly greater in the period from 1888 to 1905, and in the present period we have the fullest development of the literary ideal. Practically all of the content of reading books is now literary in a certain sense. That is, it is composed of poetry, or of fictitious stories. Only to a limited degree is factual material introduced—material intended for serious belief. History, biography, geography and science are but scantily represented.

The modern idea is that reading books must be interesting. But, what, we may ask are the qualities which will make them interesting? When Hosic examined 22 series of readers and found 4,000 titles of pieces, he discovered that 2,500 of them were named but once, that only 56 titles appeared in ten or more readers, and that only a dozen were included in two-thirds of them. May we infer that the selections appearing most frequently are the most interesting? It seems to me that we may not; and there is some evidence to support such a view.

Doctor Jordan and Doctor Dunn have lately investigated the question of children's interests in reading. The latter made up a list of 31 pieces which appeared very frequently in primary readers and which a number of adults—teachers and other school people—selected from a larger list of popular pieces as those possessing to the highest degree the quality of interest. These 31 pieces were then read, pair by pair, to school children. Something like 17,000 pupils voted as to which of a pair of these pieces they liked the better. For my own part, I am far more willing to accept the verdict of the children as to the question of interest than I am the verdict either of textbook-makers or of teachers. Now, it happened that Stevenson's "The Swing" stood at the top of the list as determined by the adults and at the bottom of it as determined by the children. Such a complete reversal is seldom found in statistical data. You recall the poem.

"How do you like to go up in a swing;
Up in the air so blue?
Oh, I do think it the pleasantest thing
Ever a child can do."

We shall have to learn that there is a difference between a poem about a child and a child's poem. There is a similar difference between prose pieces written by adults on childish topics and pieces which really possess the quality of interest for children.

It is worth while knowing what characteristics will tend most unmistakably to give a piece this desirable quality of interest. Suppose we could rate selections for interest on a ten-point scale,

and suppose we could also rate these selections for certain other qualities thought to be related to the more general quality of interest. We should then by a proper handling of our data be able to analyze interest into its constituent elements. We should be able to say, for example, whether imagery or "girlness" or "boyness" or style or realism has any bearing on the matter. This has been done to some extent by Doctor Dunn. She finds that decidedly the most important factors in producing interest for children of primary grades are a good plot and the element of surprise. Moreover, if you put an animal into the story, you add greatly to the chance of its being interesting, though it is apparently likely to be more so to boys than to girls. Moral teaching in a story seems neither to make nor mar the interest. A number of other qualities such as fancifulness, repetition, and liveliness, seem to contribute but slightly, if at all, to the quality of interest. They are correlated with it; but this appears to be more because they share with the quality of interest a common relationship to such qualities as plot, surprise, etc. Humor—at least what adults call humor—apparently contributes little or nothing to children's likes. Rhetorical style detracts slightly from interest as does the fact that a story centers about an adult rather than a child character. Familiar experiences appear lacking in interest, and realism seems most undesirable.

If then, we say that children should read what is interesting, we must ask and find an answer to the question, What is interest? and especially, What elements will give a selection this desirable quality? It is certain that adult judgments leave much to be desired. We, therefore, suggest that we first make our selection of material along the lines that have already been indicated as likely to produce interesting pieces, and second, that we check our selection by actual trial with children. In the last analysis it is they who must tell us what they should read in school.

By this I do not mean that we should not teach children interests which they do not already possess; nor do I mean to subscribe to the idea that interest is the only factor to be considered.

But whether it is interest or some idea of the good of the child, still it is in some way from the child and not from some other source that guidance must be secured.

In the present period of textbook writing it is held, as I have indicated, that reading material for children should be of literary character. As in the case of the doctrine of interest, so in this case we agree at least in part. Undoubtedly pupils should read those pieces which serve to convey to childhood and youth the cherished heritage of the race. This is not, however, altogether what those who insist on literary material have in mind. My own notion is that somewhere in the course of study of the elementary school, we should provide a high type of literary material which has an appeal for children. Sometime a great university, or perhaps a teachers' college, is going to establish a real course for the training of elementary-school teachers—a course which will include, among other things, literary criticism of reading materials for children. Such a course will do for myths and folklore and fairy tales what our graduate courses in literary criticism do for the Elizabethan dramatists and the Lake poets. It will seek to give teachers a realization of why certain materials are of worth and what their worthy qualities are. Just now we are entirely dependent upon the makers of textbooks; and I, for my part, do not believe they know much about the true characteristics of literary material for children. By the verdict of the children themselves, they know little about the interests of children.

For example, here is what happens when a textbook writer "goes after" literature. One of the most popular pieces for the seventh grade is Wordsworth's "The Daffodils."

> "I wandered lonely as a cloud
> That floats on high o'er vales and hills,
> When all at once I saw a crowd,
> A host, of golden daffodils;
> Beside the lake, beneath the trees,
> Fluttering and dancing in the breeze."

This poem occurs about as frequently as any piece in the seventh reader. I submit that a boy of twelve or thirteen who

is interested in, or who appreciates, this poem has something the matter with him. In some books the author has provided assistance in the way of questions to guide the thinking of the children. One of the textbook-makers calls these "Helps for Study," and the following are some of his questions on "The Daffodils." (1) "Tell what you can of Wordsworth's life." (2) "What figure of speech in the first line?" (This means what rhetorical figure, and the correct answer is, "A simile.") (3) "Is the figure a good one? Think of a single cloud floating in the clear sky." Further on in the poem occurs the line, "A poet could not but be gay." At this point the helps are: "Explain 'A poet could not but be gay.' Why a poet more than any other person? What is a poet?"

In short, it seems to me that in the making of readers, somewhat the same mistake has been made as has apparently been made in the case of geographies, histories, and books in other content subjects. These books are too ambitious. Indeed, perhaps it is not too much to say that nine-tenths of our textbooks are too hard for the children for whom they are intended. Doctors Franzen and Knight brought out a book a year or two ago on textbook selection. Among other things the authors found that five-sixths of the material in the typical first geography book, the book usually begun in grade IV, is too hard for the pupils of that grade to understand. They also found that only about one-half of the second book of the usual two-book series can be understood by children of grade VI—the grade in which it is first introduced.

Undoubtedly in all this we are making a mistake which differs in degree but not in kind from that of the textbook writer who within 27 pages progresses from "My son, do no ill" to "Vain persons are full of the allurements of dress." Do we not find the same thing to be true in arithmetic? Over-ambition in developing our work is preventing us from realizing the simple fundamental objectives for which the school exists and in reference to which it is judged. The knowledge and skill which is really needed in arithmetic—that is, needed in life—is

astonishingly small. Wilson found that the parents of school children had little occasion for anything beyond integers and the simplest fractions; yet we are habitually falling short of thoroughness in the fundamentals of arithmetic and spending an inordinate amount of time on percentage and business applications.

In spelling, society demands a great deal, but its demands are for a high degree of accuracy rather than for a great range of words. When we hear it said that a child cannot spell, the assertion is not made with reference to *conscientious* or *sacrilegious* or *Popocatepetl*, but with reference to *through* and *receive* and *truly*. The school is not reproached because pupils cannot name the exports of Patagonia or bound Finland or discuss the declination of the earth's axis. It is reproached because they cannot locate New York or place the Mississippi River or have a reasonable knowledge of their own state.

It is so in reading. We must ask ourselves what we are trying to do in the elementary school, and when we do so, our answer is rather definite. We are trying to teach children to read—to master that wonderful art which is such a remarkable achievement in human evolution, an artificial psycho-physical activity which, despite its artificiality, is of higher social value than any other subject or topic with which the elementary school is concerned. Now, if we ask ourselves what our more specific objectives are in the teaching of reading we shall find them fairly well represented by the things we are measuring in our reading tests—namely, rate and comprehension. If in the light of these objectives we seek an answer to our question, "What should children read in school?" we shall not look primarily for literature, unless it happens that literature is peculiarly adapted to the accomplishment of these ends. We shall take the position that in the teaching of reading, literary material by virtue of the fact that it is literary has no more place than history material or geography material or science material, and that all these types, if they are included, must possess as a basis of inclu-

sion the qualities which serve best the purposes of teaching children how to read.

Literary pieces should undoubtedly be represented and represented copiously in the program of studies, but it is a serious question whether the gems of literary expression which have formed so large a part of the matter used in the teaching of reading are really properly placed either from the point of view of teaching reading or from that of teaching literature. It seems to me that in this case our textbook-makers have tried to "kill two birds with one stone" and have dismally failed to bag either one. On the one hand, the literary masterpieces with their subtleties of thought and their mature forms of expression have not proved suitable in the teaching of reading. On the other hand, their introduction into the course of study before the child can read them, quite generally causes them to be undervalued if not actually disliked. Reports of modern investigations are replete with suggestions—and common experience is to the same effect—that the introduction of such pieces as "Evangeline," "The Daffodils," and "To a Water Fowl" have no more common effect than to arouse a distaste for them and for the types of literature which they represent. Again, in the primary grades such poems as Stevenson's "The Swing" and Sherman's "Daisies," though frequently presented and highly regarded by adults, appear to be almost despised by children.

[*To be continued*]

THE CHALLENGE OF THE PRIVATE SCHOOL

NORMAN FENTON
State Normal School, Tempe, Arizona
AND
LOWRY S. HOWARD
Whittier State School, California

The various states, according to their constitutions, assume the obligation of providing an education for their children. But this need not be interpreted as meaning that all children must be educated *in the public schools*, and in practice no such policy has prevailed. At present, indeed, the non-public schools are essential elements of the country's educational efforts. The economic value of these agencies is apparent; for by educating children they save the state money. In the private secondary and business schools alone, the population in 1917-18 was 458,000. If these private schools can be made to uphold public school standards in such matters as sanitation, attendance, personnel, and course of study, (that is minimal essentials, for to draw a rigid course of study for private schools would be to kill their most valuable possibility for experimental pedagogy), their continuance as educational agencies need not be interfered with.

That supervision, not abolition, is the growing attitude toward the private school is evidenced by the overwhelming defeat in 1920 of an amendment to abolish all non-public schools in Michigan and the passage instead of a strict supervision law. The purpose of this law is stated as follows: "It is the intent of this code that the sanitary conditions of such schools, the courses of study therein, and the quality of teachers thereof, shall be of the same standard as provided by the general school laws of the state." And thereafter follow provisions for the adequate inspection of these schools and the carrying out of the law.

In the digest of state laws relating to public education made by Hood[1] in 1915, he found that the private schools were prac-

[1] United States Bureau of Education. Bulletin, 1915, No. 47, pp. 767-770. See also Bulletin, 1920, No. 30, pp. 190-1 and Bulletin, 1922, No. 20, pp. 183-4.

tically ignored. At that time a few states—New Jersey, Illinois, Michigan, and South Carolina—required annual reports of a superficial census nature which one feels were probably never read. Since then the most significant constructive legislation, according to Hood's further studies, has been that of Nebraska and of Michigan.

Oregon recently passed a referendum which abolishes all non-public schools in the state after 1926. This act will make interesting study when it goes to the United States Supreme Court upon appeal. If the Supreme Court sustains the referendum of the citizens of Oregon, many significant developments elsewhere in the country may be expected. Since the act is not effective until 1926, it may possibly be repealed by another popular vote.

Cubberley[2] has framed an excellent set of provisions regarding the inter-action of private and public schools. In them he recognizes the state's obligation to the private schools for the service they render in educating children free of cost to the state; likewise, he recognizes the need for the imposition by the state of some limitations of freedom upon the private schools. He believes that the private schools should be given the opportunity to purchase books and supplies of the state at the same price as public schools, that those who teach in private schools should be granted certificates and permitted to attend institutes and to partake of other advantages which the public school teacher may have.

The most important aspect of non-public schools from the standpoint of the state is their effect upon the growth and ideals of the state itself. The so-called undemocratic features of the private schools have usually been brought forth as arguments against their existence in the United States. They are considered undemocratic, first, because they are exclusive, poor persons as a rule not being able to attend them; and, second, because as a community they may be snobbish and breed obnoxious ideas of superiority. There are, to be sure, all sorts of private

[2] Cubberley, Ellwood Patterson. *State and County Educational Reorganization; the Revised Constitution and School Code of the State of Osceola.* New York, The Macmillan Company, 1914. pp. 209-12.

schools, from industrial schools where boys pay a small fee and work their way, to the highly exclusive "finishing" schools. As institutions, private schools exclude the very poor either because poor boys cannot afford the tuition or because they have no means of subsistence while attending these schools. However, this objection might be equally applied to college education, even in state institutions.

Private schools should be carefully studied before judgment is passed as to the snobbery within their gates. The fraternity system of public high schools and colleges should also be reviewed with the same idea in mind. In this connection it may be observed that in military schools all boys wear the same uniform; and that girls' boarding schools often achieve the same simplicity and uniformity of dress. In public high schools, on the other hand, there is considerable variation in standards of dress and a consequent tendency to social stratification from calico to silk. There is probably more occasion for display of wealth in public schools than in the superior types of private schools. In the latter allowances are strictly limited. Pupils are not allowed to have automobiles at school with them. All outside activities are "signed up" for, not paid for directly by the students; and the number of visits to the theatre and movies is strictly limited. The superior schools permit their pupils to go to the movies once a month; or else they have pictures shown at the school one evening a week. Incidentally, the fact that these pictures can be truly censored by the school faculty is of significance.

If the graduates of the private schools were not numbered among the leaders and influential citizens, their relatively small numbers might make them less significant to the state. If, however, they train a large percentage of the statesmen and leaders of the country, then the state should find it especially to its advantage to supervise them. What are the facts regarding the quality of the population of private schools and the place their graduates take in the life of the state?

Counts[3] studied the occupations of fathers or guardians of

[3] Counts, George S. "The Population of the Private Secondary School." *School and Society*. 15:568-73. May 27, 1922.

17,263 students in the public high schools of Bridgeport, Mt. Vernon, St. Louis, and Seattle; and of 619 students in Exeter and the University of Chicago High School. The data, given in percentage form, are shown in Table I.

TABLE I. OCCUPATIONS OF PARENTS IN CERTAIN PUBLIC AND PRIVATE SCHOOLS

Parents' Occupation	Public Schools (percent)	Private Schools (percent)
Proprietors	19.8	42.7
Professional men	9.4	31.0
Managerial	16.5	11.5
Commercial	9.5	9.0
Clerical	5.8	2.1
Artisan-proprietors	4.2	1.3
Agricultural service	2.4	.7
Manual labor	29.1	.3
Occupation unknown	3.3	1.4

The first conclusion to be drawn from these data is that the private schools do not draw upon as many occupational levels as do the public schools. The public schools are more representative of the people of the country. Another feature of these statistics is that the population of private schools is more largely made up of the sons and daughters of persons of superior social status; and this also implies superior intellectual status. This last conclusion is derived not merely from inspection of the data but also from the occupation studies in connection with the army mental test data as published in the Memoirs of the National Academy of Science, Volume XV.[4]

Since the private school population, due to superior economic status, has obviously better opportunity for advanced study, the private school graduate is more likely to hold a position of importance in the community in later life than is the less favored student of the public school. This last fact, coupled with the fact that sons of professional men are more likely to become professional men than they are to enter any other type of occu-

[4] See also Dexter, Emily Smith. "The Relation between Occupation of Parent and Intelligence of Children." *School and Society*. 17:612-4. June 2, 1923.

pation, means that a much larger proportion of private than of public school students will enter professional life.

That, further, the population of private schools is composed of mentally superior pupils is shown by the figures of Table II which Howard gathered in four private schools of California in May, 1921.

TABLE II. MEDIAN SCORES IN THE TERMAN GROUP TEST OF MENTAL ABILITY

Grade	Private School "A"	Private School "B"	Private School "C"	Private School "D"	Public Schools (3 cities)	Terman Norms (1921)
6	57	...	62	...	54	...
7	79	...	98	89	82	67
8	126	...	130	125	101	87
9	127	166	129	141	118	105
10	169	176	178	172	135	122
11	151	179	...	165	155	134
12	177	186	...	179	161	148
Number of Pupils	142	64	92	47	20,000	30,000

A study of one private school, which we made in connection with Professor Terman's survey of gifted children in California, revealed the fact that there were eleven boys with intelligence quotients over 140 out of a total population of 160 pupils. The average for the public schools of the state is about one case in 200 or 250. Malherbe[5] reports a study in the Foxwood School at Flushing, New York. The average intelligence quotient was found to be 125; eight percent being over 140 and fifty percent over 125. Anderson[6] compared the scores on Army Alpha, Form 6, of 278 boys in the Hotchkiss School of Lakeville, Connecticut, with those made by high-school students of Madison, Rockford, and Sioux City—1,591 cases in all. He found that 85 percent of the Hotchkiss first-year boys exceeded 102, the median score of high-school first-year boys; that 95 percent of the Hotchkiss second-year boys exceeded 116, the median score

[5] Malherbe, E. G. "New Measurements in Private Schools." *The Survey.* 46:272-3. May 28, 1921.

[6] Anderson, John E. "The Intelligence of a Highly Selected Group." *School and Society.* 16:723-5. December 23, 1923.

of high-school second-year boys; that 88 percent of the Hotchkiss third-year boys exceeded 126, the median score of high-school third-year boys; and that 85 percent of the Hotchkiss fourth-year boys exceeded 134, the median score of high-school fourth-year boys.

The fact that the pupils taught in such private schools are highly selected is exactly the point. What we wish to show is, first, that these schools contain superior mentality, in order that we may add the second point, namely that, so far as the school authorities of California or New York or Connecticut are concerned, these boys, for the most part sons of citizens of these states, might be studying in China or the South Sea Islands. Non-public elementary schools are never inspected. The high schools may be given a genteel visit once yearly (California) to ascertain whether the schools are deserving of certification. We believe it to be essential that the state should know all about the private schools within its borders. It must protect itself and the well-being of future citizens whom it now leaves to the educational devices of corporations or private individuals.

Such facts as we have stated afford sufficient reason for the state to interest itself in the private schools. What better reason could there be for this than the fact that the pupils of these schools are in many cases the future leaders of the country? For, in addition to their superior intelligence, they have the means and the inclination for higher educational opportunities. These facts make apparent the need for adequate supervision and direction by public school officials of the work of the private schools.

Above all, perhaps, the private schools, as Professor Cubberley has said recently, are an eternal challenge to the public schools. The growth of private schools in a community often indicates that the public schools are on the wane or in politics. Private country day schools and their success—what are they but challenges to the state to provide similar advantages for its children? Private 24-hour schools—what are they but challenges to the state also to provide such schools? Since the private schools are drawing many of the future leaders of the state,

what is this but a challenge to the state to train its own leaders by providing better 24-hour institutions and country day schools than do private agencies? At present the state provides 24-hour schools for imbeciles, degenerates, and delinquents. Why not for normal and superior children? Therein is the challenge of the private school, and whether or not the state accepts this challenge is a test of its foresight.

A good deal of the past history of American education involves just such acceptances of challenges from the private schools. In fact, the history of education in the past centuries is a recital of the growth of the public school system, of its taking over from private agencies the work of educating children and adults. Will the state now accept these two challenges of the private schools, namely, 24-hour schools and country day schools?

One might go on in optimism and see state 24-hour schools for superiors and a national university for superiors! Entrance to these might be based upon ability; and residence thereat might be free, or with some slight pay arrangement as at West Point. These boys and girls would thus be given the sense of state service all the time they were at their studies. Indeed, a term of service to the state might perhaps be required of them all at the completion of their course.

The Catholic Church is reported to be planning a one million dollar normal school to be built in San Francisco. It will turn out teachers, trained in the methods of the state normal schools and teachers' colleges, who have taken the vow to devote their lives to teaching and who are actuated by religious enthusiasm. This is another challenge to the public school of the future—a challenge to the state school system to turn out teachers whose devotion to the cause of education will exceed in fervor that of the parochial school teachers, and whose achievements in teaching will excel in measurable educational product. To accomplish this, some spiritual attachment to the ideals of the public schools similar in its working to the religious motives of the teaching nun will have to be inculcated in our teachers. This is, indeed, the greatest challenge of all.

WHAT IS MEASURED BY INTELLIGENCE TESTS?

OMEN BISHOP
State University of Iowa

During the last two years so much discussion as to just what is measured by the group intelligence tests has taken place, and there have been so many published expressions of opinion concerning this important problem, that it has seemed wise to undertake some experimental study of the question. That there are essential individual differences in native endowment among the pupils in our schools as well as among ordinary people in the whole population, is probably not open to question. Such differences are so noticeable that no competent worker in the field of education would doubt their existence, even without the aid of standard tests. There has been, however, a rather naïve assumption current among the active promoters of group intelligence testing that the inborn raw brain power accounted for such a large proportion of such individual differences that other factors might be deemed negligible. Although some have qualified their claims upon this subject both in published literature and in addresses, still the general impression gained from the perusal of the literature on intelligence testing is that one of these ten-page tests—or maybe two of them for safety—if applied according to directions will enable one to tell with scientific accuracy the native inborn mental fibre of both groups and individuals.

That these instruments do not test or measure this native gift directly is admitted; but it is believed that they measure indirectly by measuring the results of intelligence. The conclusion is then reached that the results are proportional to the amount of ability which produced the results. This position is doubtless tenable if all causes are considered which bear upon the results. It is just here, however, that the critical technique of statistical method must be adhered to most carefully. The principle has long been accepted that where there are a number

of variables operating to produce a result, the influence of any one of them should be determined by reducing all to constants except the one to be measured. In other words, all factors, except the one which is under consideration, should be eliminated or allowed for.

Now it is the practice of most of our educational workers to teach material very similar to most of the items found in group intelligence tests. Accordingly, the assumption is made that all who take the tests have had approximately equal opportunity to learn, both in school and in their general environment, and that when this factor is assumed to be equal in all, it may be deemed negligible. Although these advocates acknowledge that one school or one teacher may be better than another, and that the home environments of the pupils may not be equal, yet they believe that by and large these factors produce about the same results among pupils of equal intelligence. This common assumption in intelligence testing will appear more emphatically as this article proceeds.

The study which is here reported was undertaken to determine by experiment, at least in part, what effect the teaching of certain things would have upon the scores in group intelligence tests. Specifically, this investigation was an attempt to show whether modified teaching under ordinary school conditions would result in modified intelligence test scores.

The test selected for use throughout the experiment was the Otis Group Intelligence Scale, Form A. This choice was influenced by several considerations: first, the general educational literature seemed to indicate that it is as widely used as any test; second, a study by Breed and Breslich[1] within the last two years has shown that the Otis scale has a closer correlation with the results obtained by using a number of tests than any other one of them when used alone; third, the author of the Otis test has put forth certain claims which bring the problem of just what is tested to a definite statement. In the

[1]Breed, F. S. and Breslich, E. R. "Intelligence Tests and the Classification of Pupils." *School Review*. 30:51-66; 210-26. January and March, 1922.

introduction to the "Manual of Directions" issued by the author to accompany the test, Doctor Lewis M. Terman, himself a pioneer in the field of intelligence testing says: "The Otis Group Intelligence Scale is an instrument for the measurement of native mental ability. It is applicable to any individual whether child or adult, who has had the equivalent of three or four years of schooling. With subjects of this much schooling, the Otis scale probably comes as near testing raw brain power as any system of tests as yet devised."

Doctor Otis also refers in his Preface to Doctor Terman's work as follows: "In order to attain a comprehensive insight into the purposes and uses of intelligence tests, one cannot do better than read the first chapter or two of Doctor Terman's book, *The Measure of Intelligence*." These chapters are clearly permeated with the idea that it is native brain power which in the final analysis is being tested. With a claim that the Otis test is as good as any other, the present writer has no argument, and the mutual exchange of amenities is most pleasing; but as one reads further he constantly meets the assumption that not only is the instrument in question as good as any other but that it may be used with confidence that the scores obtained give a correct statistical measure of the individual's raw brain power. This study was made for the purpose of contributing in an experimental way to the problem of determining whether this assumption and the consequent conclusions are justifiable.

The experiments here recorded were carried out at the University High School, State University of Iowa, in March and April, 1922. The pupils were of four groups, the first and second groups being from the seventh and eighth grades of the junior high school, and the third and fourth groups from the ninth and tenth grades. The following procedure was used: These pupils were given the Otis Group Scale, Form A. Upon the showings of this test they were divided very carefully into four groups, two of seventh- and eighth-grade pupils, known in this report as groups A and B, and two of ninth- and tenth-grade pupils, called groups C and D.

In making this division, every child in group A was paired with some child in group B. Attention was paid not only to the Otis scores, but also to age, sex, apparent physiological development, and habits of work as shown by class performance. It was of course impossible to make these pairings equal in all these particulars; but great care was used to make them as nearly equal as possible. If a number of pupils had Otis scores which varied by only a small amount, the plan adopted was to pair two large strong pupils, or two small rather under-age pupils, even if it would have been possible to make the pairings a little more equal on the basis of the Otis scores alone. It was thought that pairing a large mature boy with a small precocious lad who had been thoroughly school trained was hardly just procedure even though their original scores were very close. Similar pairings were made in groups C and D. In all, 64 children were used in this study.

A series of ten lessons was then prepared by the writer, each very similar to one of the ten pages of the Otis test. These lesson tests were used deliberately to teach the principles involved in the Otis test without teaching the exact materials of the scale. In general, the wording was changed, and the problems, vocabularies, and situations were modified so that the responses learned in the teaching could not be carried over to the original test. In vocabularies the words were taught, but with changed context. Similar arithmetical problems and geometrical figures were introduced which were to teach the principles used in the Otis test. In this way each page in the Otis scale was subjected to a process which formed part of an attempt to teach its elements under quite ordinary school conditions. Ordinary supervised class study was employed.

It would be obviously impossible for any teaching to make amends in a few short lessons for all the faults of years of loss from improper teaching and environment, in all fields. The work was, therefore, confined very closely to the immediate principles, vocabularies, processes, etc., which were necessary for the work in hand. This is admitted and carries its own

justification. Nothing was introduced which any pupil who had been reared in a favorable surrounding and well taught in home and school might not have learned in the ordinary course of events; but there are some things which the child who has not been reared in such an environment might not have learned. In the class instruction, the lesson tests were often changed from day to day, so that the pupils might learn the principles involved and not memorize the answers.

The first five lessons were taught to group A, and the last five lessons, to group B; five of the lessons were, therefore, omitted with each group. The lesson materials were first given as a test to each group, then the teaching process was employed. This was in turn followed by the re-administering of the materials in the test form. The purpose of this technique was to show statistically that the materials had actually been taught to the groups with some degree of success. The time used in teaching the lessons varied greatly among the different pupils, owing to the absence of some from the classes on account of sickness, the demands of the regular school work, and the fact that some soon mastered the lessons and could no longer be interested. The median amount of time was approximately 150 minutes for each pupil in the group. This teaching and testing was followed by the repetition of the Otis Scale, Form A, in its complete form, after which the results were tabulated. Then the same procedure was used for group B. The results here were not quite so marked, as it was discovered that they had been interested in what their friends in group A were doing and had talked over much of the matter in spite of all the precautions that were used.

Again the same method was used for group C, except that all ten pages of the lessons were taught instead of just five. More time was used for this group than for A or B since they had twice as much to learn. To group D nothing was taught. The Otis scale was simply repeated to show for a check group how much the second taking of the test increased the scores.

Results

In group A, which studied the first half of the lesson tests, there was a total gain in scores of 376 points on the first five pages of the Otis scale, as against 42 points on the last half where they had not studied similar material. In percentage form individual gains upon the first five tests of the Otis scale had a range from 26 percent to 140 percent, with a median of 40 percent. On the second half of the test, that was repeated without study, the percents ranged from —10 to 21 with a median of 6. This shows a net gain of 34 percent on the part which they had been taught.

In group B in which the pupils were not taught the first half, but were taught the second, the range of gains on the taught half was from 15 to 85 percent, median 31 percent, as compared with a range of 4 to 39 percent, median 15 percent, upon the untaught portion. Figure 1 gives the initial and final median scores made by groups A and B.

In comparing groups C and D, which were also made approximately equal in the pairings, the following results were obtained: group C, the pupils of which were taught all ten pages of the test, made a total gain of 672 points as against a total gain of 243 points by the untaught group D. The range of individual gains in percentage form for group C was from 18 to 57 percent with a median of 30 percent. This is to be compared with group D, whose range of gains was from 1 to 39 percent, with a median of 11 percent. The report of the original and final testings with the percentage gains of each pupil are given in Table I.

There were many interesting things which appeared during the study, such as evidence that the pages are not of equal difficulty to learn, and correlations upon the elements of the test. These must, however, be omitted from the present report.

Conclusions

The results of the study seem to indicate that the scores on this group intelligence scale are very greatly influenced by the

teaching the pupils have received. In every group and in almost every individual case the evidence is clear. Group A, the members of which were taught material similar to the first half of

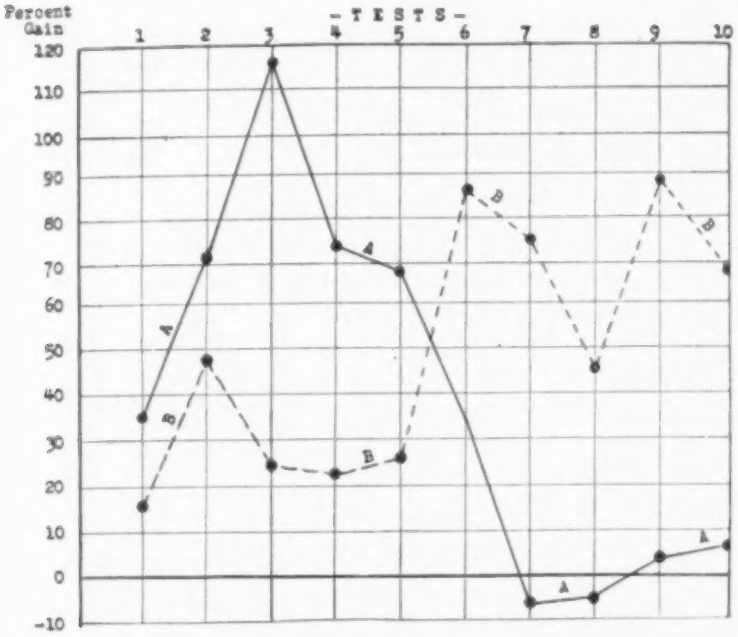

FIGURE 1. PERCENTS OF GAIN MADE BY GROUPS A AND B ON EACH TEST

the test, made a greatly improved score upon that half and only a slightly improved score upon the other half. The fact that a pupil can double his score on an intelligence test as the result of a few lessons upon similar material—not identical material—makes the teaching factor seem particularly important. Something other than native intelligence seems certainly to be in evidence. Group B, which was taught material similar to the second half, gained greatly on that half, and gained but little on the first half. This was in spite of factors, already mentioned, which operated to make them equal.

Group C, having had instruction on material similar to all ten pages, made large gains on all ten parts, while group D, with no teaching except regular school work, made gains about a third as great. More striking still are some of the results as shown in individual cases. The writer regrets his inability to state objectively some of the information bearing upon these

TABLE I. SHOWING OTIS SCORES AND GAIN PERCENT GROUPS C AND D

Pupil	First Testing		Second Testing		Percentage Gains	
	C	D	C	D	C	D
1	185	190	218	197	18	4
2	178	173	210	193	18	12
3	169	169	200	192	18	14
4	167	168	209	182	25	8
5	165	167	196	169	19	1
6	163	165	202	189	24	15
7	156	151	196	152	25	1
8	147	148	186	159	26	7
9	138	142	186	161	35	13
10	138	140	186	156	35	11
11	127	121	178	127	40	5
12	116	117	161	135	39	15
13	111	111	144	119	30	7
14	109	109	154	127	41	17
15	107	107	154	127	44	17
16	99	102	128	113	30	11
17	68	64	107	89	57	39

Median Gain percent of Group C—30
Median Gain percent of Group D—11

cases. He gathered the data as best he might from reports of previous training and of general home environment. He used them with caution because he knows of no way to measure objectively the loss suffered by a child who has been kept out of school much of the time, or how to compare objectively the influence of a low home environment with that of a home of culture. Such social factors were very noticeable in the cases of pupils reported as Numbers 2, 7, 14, and 15 in group A, and 15 in group B. These pupils made greatly improved scores on the second test and their original low scores seem to have been produced largely by unfavorable early teaching and environment.

The scores of these pupils on the two tests are graphically shown in Figure 2. It seems, therefore, that the use of a group intelligence test to determine raw brain power and the interpretation of the results as showing definitely whether an individual is capable of learning when given good normal environment, is liable to great error. There are other factors such as teaching, home, and associates which must be included. Any interpreta-

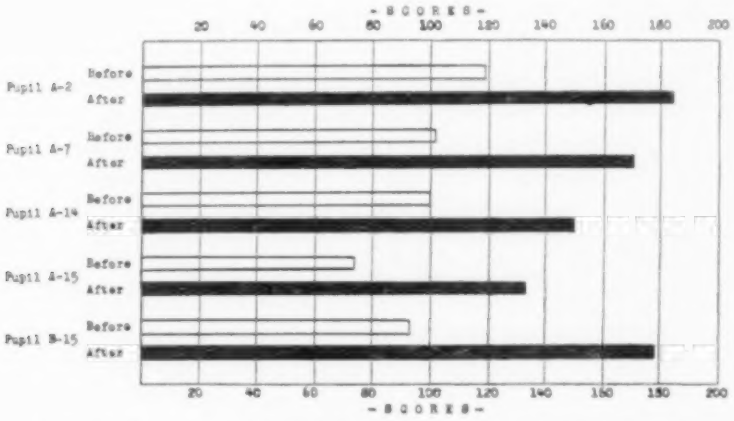

FIGURE 2. CHARACTERISTIC INDIVIDUAL SCORES BEFORE AND AFTER STUDY

tion of scores which treats these influences as negligible is not justified. In fact, the common practice of calling these tests intelligence tests, will in many cases involve most serious error if by intelligence is meant native ability to learn under standard conditions and opportunities. If by the term intelligence is meant a composite of qualities which are partly native and partly the result of environment, then the literature which speaks of it as inborn must be discarded. As an instrument for measuring the ability to do school work the tests doubtless are of value, since in much of our school work the ability to succeed depends not only upon original bonds and the ease with which new ones may be formed but also upon those bonds which have actually been formed through the learning process and which are func-

tioning in skills and habits. Whether with our present knowledge it is possible to construct a group test which will be independent of every factor except pure ability of the kind usually meant when a child is spoken of as naturally stupid, or naturally bright, or normal, we cannot say. At least, until we shall have achieved that much-to-be-desired end, it would be well not to give the impression to an eager public that we have accomplished it.

THE SPEED FACTOR IN MENTAL MEASUREMENTS

GILES MURREL RUCH
State University of Iowa

The attacks on the validity of modern tests of general intelligence by William C. Bagley, Walter M. Lippmann, and others, have raised a number of important issues. One of these is the question of the effect of time limits on the scores of subjects taking such tests. Mr. Lippmann has stated that the intelligence rating earned by a soldier was chiefly determined by the time limits used in giving the tests. If Mr. Lippmann had in mind, as would seem to be the case, that intelligence scores earned under regular time limits would be surpassed if added time allowances were given, there is probably no psychologist who would not grant the truth of his somewhat naïve contention. If this is not true, it constitutes the first exception in the literature of the improvability of mental functions. What must not be done, however, is to confuse this issue with the separate question of whether general mental ability (in the sense involved in the definition of *mental age* or other units of mental measurement) is raised by increased time limits. *Scores in intelligence tests or educational tests possess meaning only in reference to the norms of performance established for the particular test.* It is surprising that this elementary concept in mental measurement should have occasioned confusion even in the mind of the layman. Complete removal of time limits would probably invalidate the norms of most of the existing tests but it is unlikely that anything more serious would happen in such an event. The norms might easily be reestablished upon a *power* basis.

There is an aspect of this general problem, however, which does need further study. Is it not likely that certain individuals who react slowly in the test situation are to some degree penalized and underrated by "speed" tests? Even if "speed" and "power" are highly correlated in most or all mental functions, this relationship, for many reasons, is not perfect. Such cor-

relations are never unity. Language difficulty often makes for error. The reaction times of individuals differ as greatly, presumably, in the higher mental processes as in the simple reaction experiments. Some subjects, as the result of volitional, intellectual, or scholastic habits, take the test situation more seriously than do others. Perhaps both the reliability and the validity of most tests and measuring scales would be improved by re-standardization upon a complete "power" basis. At any rate, the way to the proof or disproof of this contention is obvious.

The reasons for time limits are chiefly practical ones concerned with economy of time. If the time limits of all tests could be experimentally established in such a way that from 90 to 95 per-cent of the subjects could complete all of the items within their abilities in the allotted time, the practical situation would appear to be met satisfactorily. It must be admitted that many of the existing tests do not meet this criterion. In correlational terms, if the correlation between scores earned under regular time limits and those earned in unlimited times is found to be as high as 0.95 to 0.99, especially if they are as high as or higher than 0.98, there would be little practical justification for greatly increased time allowances.

The argument for this conclusion can be stated very nicely by the use of Kelley's "Coefficient of Alienation" based upon the standard error of estimate. If $\sigma_{1.2}$ is the standard error of estimate of a first variable from known values of a second variable, σ_1 the standard deviation of the first variable, and r the coefficient of correlation between the two variables, then

$$\sigma_{1.2} = \sigma_1 \sqrt{1-r^2}.$$

The factor $\sqrt{1-r^2}$ is the coefficient of alienation. According to this formula, if the correlation between scores under regular time limits and those under no time limits were zero, the standard error of estimate of the latter from the former would equal the standard deviation of the distribution of scores earned under regular time limits. In other words, $\sigma_{1.2}$ would equal σ_1, since the factor $\sqrt{1-r^2}$, or the coefficient of alienation, would be 1. Table I shows the alienation of coefficients of correlation away

TABLE I. COEFFICIENTS OF ALIENATION CORRESPONDING TO HIGH VALUES OF THE COEFFICIENT OF CORRELATION

(a) Coefficient of Correlation r	(b) Coefficient of Alienation	(c) % of Decrease in Standard Error of Estimate over 0.000 Correlation
0.950	0.3122	68.78
0.960	0.2800	72.00
0.970	0.2431	75.69
0.980	0.1990	80.10
0.990	0.1411	85.89
1.000	1.0000	100.00

from unity in terms of the amount of reduction of the standard error of estimate for values of r from 0.95 to 1.00. To take a concrete illustration, the standard deviation of the scores for the Terman Group Test (see Table III below) is 30.7 for the total 150 cases under regular time limits. The coefficient of correlation between scores with regular time limits and scores with unlimited time is 0.960 (see Table II). Therefore the standard error of estimation of scores for unlimited time from regular scores is about 8.6 units. The value 8.6 was obtained by multiplying 30.7 by the value 0.2800 found in column (b) opposite 0.960 in column (a). Had r been 0.98, $\sigma_{1.2}$ would have been 6.1 units, and had r been 0.99, $\sigma_{1.2}$ would have been reduced to 4.3 units.

It follows that if the correlation of regular time with unlimited time falls below 0.95, the "speed" element does begin to be serious. At $r = 0.866$, the error of estimate becomes equal to one-half of the standard deviation of the distribution of regular scores, an error which is of great practical significance.

Review of Previous Studies on the "Speed" Factor in Tests

It seems that only two serious efforts at the study of power *vs.* speed have been made. The first is the experiment carried

out by May and Terman[1] during the war with Army Alpha. Five hundred and ten soldiers were given Army Alpha under regular and under double time limits. The correlation between the two sets of scores was 0.967. The conclusion drawn was that increasing the time limits made no material differences in the rating of the soldiers. The second investigation was that of Ruch and Koerth[2] with the same test, using college students. The correlation between single time and double time for 122 college students was found to be 0.966, and for single time with unlimited time to be 0.945 (the correlation ratios were about 0.955, giving evidence of slight curvilinearity). In the case of the second study some of the subjects worked four or five times the regular time limits and the results were as nearly power scores as was possible to obtain. Admitting that Army Alpha is a speed test to a considerable extent, no great injustices are likely to be involved in its administration under standard conditions.

The Present Investigations

It was thought to be very desirable that the types of studies reported above be extended to other types of tests, both mental and educational. Five further minor experiments have been carried out, three dealing with the Terman Group Test and two dealing with educational tests. These last include the reading and arithmetic tests of the Stanford Achievement Test.[3]

Terman Group Test.—The Terman Group Test of intelligence was given in the seventh and eighth grades of the schools of Rock Island, Dundee, and Villisca, Iowa. The method used was similar to that reported by Ruch and Koerth in that the pupils first took the tests under standard conditions using black-leaded pencils. These pencils were then collected and red-leaded

[1] Yerkes, Robert M. ed. *Psychological Examining in the United States Army.* Washington, Government Printing Office, 1921. p. 416. (Memoirs of the National Academy of Sciences, Vol. 15.)

[2] Ruch, G. M. and Koerth, Wilhelmine. "Power vs. Speed in Army Alpha." *Journal of Educational Psychology.* 14:193-208. April, 1923.

[3] Kelley, Truman L., Ruch, Giles M., and Terman, Lewis M. *Stanford Achievement Test.* Yonkers-On-Hudson, New York, World Book Company, 1922.

pencils were substituted. The pupils were then allowed to work as long as they wished with the red pencils before they surrendered their booklets to the examiner. Since they were not allowed to make any erasures, it was easily possible to keep the scores earned in the two periods separate. A record was kept of the total additional time used during the red pencil period. Table II presents the correlation between single time and unlimited time for each of the three schools.

TABLE II. CORRELATIONS BETWEEN SINGLE TIME AND UNLIMITED TIME FOR THE TERMAN GROUP TEST

City	r	P. E.	N
Rock Island, Iowa	0.937	0.016	28
Dundee, Iowa	0.955	0.010	35
Villisca, Iowa	0.970	0.004	87
Three schools pooled	0.960	0.004	150

TABLE III. THE MEANS, STANDARD DEVIATIONS, AND COEFFICIENTS OF VARIATION FOR THE THREE SCHOOLS, SINGLE AND UNLIMITED TIMES

City	Means		Standard Deviations		Coefficients of Variations	
	Single	Unlimited	Single	Unlimited	Single	Unlimited
Rock Island, Iowa	81.6	96.7	31.9	34.8	.391	.360
Dundee, Iowa	79.4	95.4	35.6	35.8	.448	.376
Villisca, Iowa	83.2	99.4	28.4	29.6	.341	.298
Three schools pooled	82.1	97.9	30.7	32.1	.374	.328

Tables II and III give most of the significant facts about the effect of time limits in the Terman Group Test. For ranges of talent represented by a standard deviation of approximately thirty score units, the correlation between standard time and unlimited time is about 0.960 ± 0.004. The average pupil gained almost fifteen score points. The standard deviations grew slightly larger under the increased time, but variability in the sense of the Pearson coefficient of variation appeared to decrease to some extent. The probable error of estimation of scores un-

der unlimited time from those earned under regular limits was 5.8 score points (pooled data).

Educational tests.—The reading and arithmetic tests of the Stanford Achievement Test were given to a group of 64 sixth-, seventh-, and eighth-grade pupils in the University Elementary School at Iowa City, Iowa. The range of talent of this group is represented by the fact that the standard deviation of the mental ages of these 64 pupils is about 22 months. The plan of colored pencils for registering the work done in the extra-time period was followed exactly as before. It should be pointed out in connection with the Stanford Achievement Test that the time limits in these tests are not sharply drawn but are determined in such a way that perhaps 90 percent of all pupils can finish all the items in the time allowed. The reading test consists of three separate tests, viz., word knowledge, sentence meaning, and paragraph meaning. The arithmetic test is composed of a test of computation and a test of reasoning problems. The total working time for the three reading tests is 40 minutes and the total working time for the two arithmetic tests is also 40 minutes. For unselected age groups the reliability of a single form of these tests is about 0.95 for reading and 0.90 for arithmetic. As before, every effort was made to have each pupil "work himself out" before surrendering his paper to the examiner. Very few pupils worked more than double time in the reading and arithmetic tests. A considerable number failed to gain even a single point by the additional time allowed.

Table IV shows the correlations between single and unlimited time scores in both educational tests. Table V presents the means, standard deviations, and coefficients of variation.

TABLE IV. CORRELATIONS BETWEEN SCORES EARNED IN SINGLE AND UNLIMITED TIMES FOR THE READING AND ARITHMETIC TESTS OF THE STANFORD ACHIEVEMENT TEST

Test	r	P. E.	N
Composite reading.....	0.968	.005	64
Composite arithmetic....	0.976	.004	64

TABLE V. THE MEANS, STANDARD DEVIATIONS, AND COEFFICIENTS OF VARIATION FOR THE READING AND ARITHMETIC TESTS OF THE STANFORD ACHIEVEMENT TEST, SINGLE TIME AND UNLIMITED TIME

Test	Means		Standard Deviations		Coefficients of Variations	
	Single	Unlimited	Single	Unlimited	Single	Unlimited
Composite reading...	195.9	205.4	37.2	34.1	.190	166
Composite arithmetic	199.8	205.9	40.1	45.7	201	222

Inspection of Tables IV and V shows very clearly that the time limits used in the Stanford Achievement Tests do not invalidate the tests since they are, for all practical purposes, *power* tests as administered. The correlations reported for single *vs.* unlimited time probably approach unity about as closely as the reliabilities of the tests permit.

Conclusions

1. The scores on the Terman Group Test earned under regular time limits correlate as high as 0.960 for ranges of talent roughly equal to combined seventh and eighth grades.

2. The scores earned in the reading and arithmetic tests of the Stanford Achievement Test under regular time limits correlate as high as 0.968 and 0.976, respectively, for ranges of talent equal to that represented by a standard deviation of mental ages of about 22 months.

3. In tests such as these studied, the speed factor does not seriously influence the ratings of the pupils.

4. It is suggested that a good working rule for the establishment of time limits, where pure power conditions are not practicable, would be to set the time limits so that 90 to 95 percent of all pupils can finish every item within their abilities.

A STUDY OF INTELLIGENCE SCALES FOR GRADES TWO AND THREE

J. Cayce Morrison, W. B. Cornell, and Ethel Cornell
The State Department of Education, Albany, New York

Teachers and supervisory officers are continually seeking information concerning the group intelligence tests which will give most satisfactory results. The study described in the following pages was undertaken to determine, if possible, which of several tests would prove most useful in grades II and III.

Six tests were given: Detroit First-Grade Intelligence Test; Haggerty Intelligence Test, Delta I; Dearborn Group Tests of Intelligence, Series I, Revised Edition; Pressey Primer Scale; Otis Group Intelligence Scale, Primary Examination; and Myers Mental Measure.

Administration of the Tests

Two classes were tested with each of the six tests. One was a third grade of 45 pupils; the other, a second grade of 39 pupils. The tests were given in the forenoon of three different days during the week, January 18-25. The Detroit and Haggerty tests were given the first day; Dearborn and Pressey, the second; Myers and Otis, the third. The writers took turns in giving the tests; but in each case where one gave the test, another checked the timing and procedure. Each paper was scored by two different examiners and all differences in scoring were adjusted through conferences.

The Data

Table I gives the distribution of scores, the total number of pupils tested with each test, the median score, the average deviation from the median, the coefficient of variability, the standard deviation, and the probable error of the median. Since there were but two grades tested, Roman numerals II and III are used in all tables to refer to grade II and grade III, respectively.

TEST I. DISTRIBUTION OF SCORES FOR EACH CLASS ON EACH OF THE SIX TESTS

Score	Dearborn		Detroit		Haggerty		Myers		Otis		Pressey	
	II	III	II	III	II	III	II	III	II	III	II	III
150		1										
145		3										
140		2										
135		4										
130		2										
125	2	1										
120	1	1										
115	2	6										
110	2	2										
105	3	3										
100	1	1										
95	4	6										
90	3	2										3
85	2	2				1					5	2
80	2	1										
						1				2	5	7
75	6	3				3				4	7	6
70	1	1									6	6
65		2				3					6	5
60	6				2	6		1		2	6	5
55	2				3	6		2	2	4	3	3
						2			5	5		1
50			1									
45			3	4	1	6		1	8	7	5	6
40			11	13	6	5	2	6	8	6	1	2
35			17	22	9	4	3	1	5	6		1
30			6	5	6	3	7	9	3	5		
25					5	2	6	10	7	2		1
20					3		6	7	1		1	
15			1		3	1	8	4				
10					1	1	4	2				
5						1	1					
Total	37	43	39	44	39	44	37	43	39	43	39	43
Median	88.7	111.3	38.7	38.9	35.8	49.6	24.6	29.3	42.2	46.8	67.9	67.1
A. D.	17.6	20.2	4.1	3.3	9.3	13.8	7.6	8.6	7.7	10.7	10.0	11.7
C. V.	.198	.182	.106	.083	.261	.277	.310	.294	.183	.230	.148	.174
S. D.	20.6	23.8	5.74	4.0	12.13	17.02	8.9	11.38	9.4	13.5	13.2	14.4
P. E.	2.86	3.0	.777	.509	1.64	2.169	1.24	1.467	1.273	1.75	1.871	1.861

Criteria for Rating the Tests

Various criteria have been proposed for evaluating group tests.[1] The ones employed in this study are: (1) correlation with the composite of the six scores, (2) correlation with the mental age as determined by the Stanford Revision of the Binet Test, (3) the discriminative capacity of the six tests, and (4) the administrative features of the tests.

[1]Henmon and Streitz. "A Comparative Study of Four Group Scales for Primary Grades." *Journal of Educational Research.* 5:189. March, 1922.

The Composite Score

The average score and standard deviation (σ) was calculated for each class on each of the six tests. Then, each pupil's score on each test was reduced to a sigma (σ) deviation from the average. The sigma deviation was in turn reduced to a point scale with 50 as the central tendency and one point of the scale equal to one-tenth sigma. The following data for pupil Number 1 of grade II illustrate the method:

	Dearborn	Detroit	Haggerty	Myers	Otis	Pressey	Composite
Test Score	64	44	17	20	28	55	.....
d σ	−1.15	+0.86	+1.46	− .79	−1.31	−1.08	.....
Calculated Score	38.5	58.6	35.4	42.1	36.9	39.2	41.9

Table II shows the correlation of each test with this composite score.[2]

TABLE II. CORRELATION OF EACH GROUP INTELLIGENCE TEST WITH THE COMPOSITE SCORES

Group Test	Correlations		Average of Two Grades	Rank
	II	III		
Dearborn	.862	.824	.843	2
Detroit	.763	.852	.808	3
Haggerty	.775	.823	.799	4
Myers	.832	.690	.761	5
Otis	.868	.909	.889	1
Pressey	.752	.706	.729	6

The Mental Age

During the five weeks immediately following the group testing, each child was tested individually with the Stanford Revision of the Binet Scale.[3] While the Binet scale may not be

[2] All correlations were calculated by the product-moment formula. The correlations in Tables II and III were calculated by two different methods: (a) Toops Correlation Charts, and (b) the method given by Rugg in his *Statistical Methods Applied to Education*, pp. 251-76. Wherever the difference between the two results exceeded two-hundredths (.02), corrections were made.

[3] The individual testing was done by the Doctors Cornell assisted by Miss Rachel Dunn and Miss Justine Devlin of the Albany public schools. Dr. Ethel Cornell checked all records and prepared the report.

a perfect test, it is usually conceded to be the best single measure of intelligence or of ability to do school work that has yet been devised for use in primary grades. The cost of administering it, however, has kept it from coming into general use. Since this is true, there is distinct value in finding which group test comes nearest to serving the same purpose. Table III gives for each class the correlation of each group test with the mental age as determined by the Stanford-Binet test.

TABLE III. CORRELATION OF EACH GROUP INTELLIGENCE TEST WITH THE MENTAL AGE DETERMINED BY THE STANFORD REVISION OF THE BINET SCALE

Group Test	Correlations		Average of Two Grades	Rank
	II	III		
Dearborn	.544	.490	.517	4
Detroit	.544	.572	.558	2
Haggerty	.653	.447	.550	3
Myers	.400	.575	.488	6
Otis	.623	.696	.660	1
Pressey	.530	.494	.512	5

On the basis of the correlation with the mental age as determined by the Binet scale, the group tests ranked in order as follows: Otis, Detroit, Haggerty, Dearborn, Pressey, and Myers. When the coefficients for the two grades were considered separately, the Otis test ranked first for the third grade and second for the second grade. The Haggerty test ranked first for the second grade and sixth for the third; while the Myers changed from rank six in the second grade to rank two in the third. Again considering the average of the two coefficients, the Otis test was easily first, the Haggerty and Detroit had coefficients almost equivalent; and the Dearborn and Pressey very nearly tied for fourth place.

Discriminative Capacity

Henmon[4] and Holley[5] have both used discriminative capacity as a criterion for judging tests. Henmon says: "Other things

[4] Henmon and Streitz. *Loc. cit.* 5:185-94. March, 1922.

[5] Holley, C. A. *Mental Tests for School Use.* Urbana, Illinois, University of Illinois, 1920. pp. 1-91. (University of Illinois Bulletin, v. 17, no. 28, March, 1920, Bureau of Educational Research Bulletin, no. 4.)

being equal, that test is most useful which segregates most clearly and unambiguously different age or grade groups. The larger the difference between the averages or medians of six-year-olds and seven-year-olds or between first-grade and second-grade children, the better the test for classificatory purposes. More important than the difference between the central tendencies of two age or grade groups is the variability in each group. The smaller the coefficient of variation or the smaller the percent of overlapping between successive years, the greater the discriminative capacity and usefulness of the test. For practical purposes in these respects a comparison of grade groups is more useful than a comparison of age groups."

Table IV gives the data and ranking for the six tests according to their discriminative capacity based on the coefficient of variability.

TABLE IV. SHOWING HOW THE SIX TESTS COMPARE IN DISCRIMINATIVE CAPACITY

Group Test	Difference between Medians	P. E. of Each Difference	Coefficient of Variation	Rank
Dearborn	22.5	4.20	0.186	1
Detroit	0.19	.93	4.894	5
Haggerty	13.72	2.72	0.198	2
Myers	4.67	1.92	0.411	3
Otis	4.57	2.16	0.472	4
Pressey	−.85	2.57	−3.023	6

The percent of second-grade children who exceeded the third-grade median on each test was: Dearborn, 17.6; Detroit, 48.4; Haggerty, 12.8; Myers, 34.9; Otis, 31.3; and Pressey, 52.7. The percent of third-grade children who fell below the second-grade median was: Dearborn, 19.8; Detroit, 48.1; Haggerty, 19.9; Myers, 28.9; Otis, 36.3; and Pressey, 52.3. The average of the two percents for each test was: Dearborn, 19; Detroit, 48; Haggerty, 16; Myers, 32; Otis, 34; and Pressey, 52. These percents of overlapping are illustrated in Figures 1 to 6.

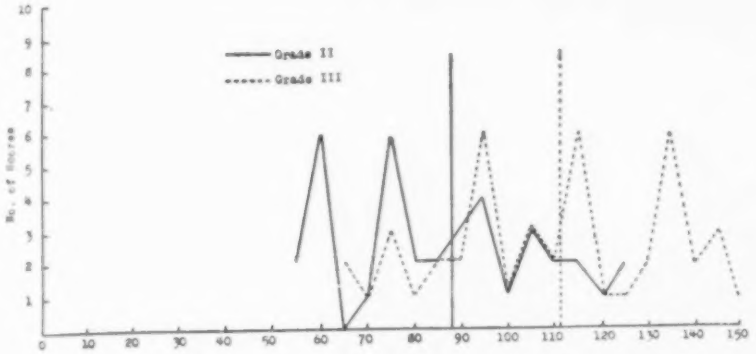

FIGURE 1. DISTRIBUTION OF SCORES ON THE DEARBORN TEST

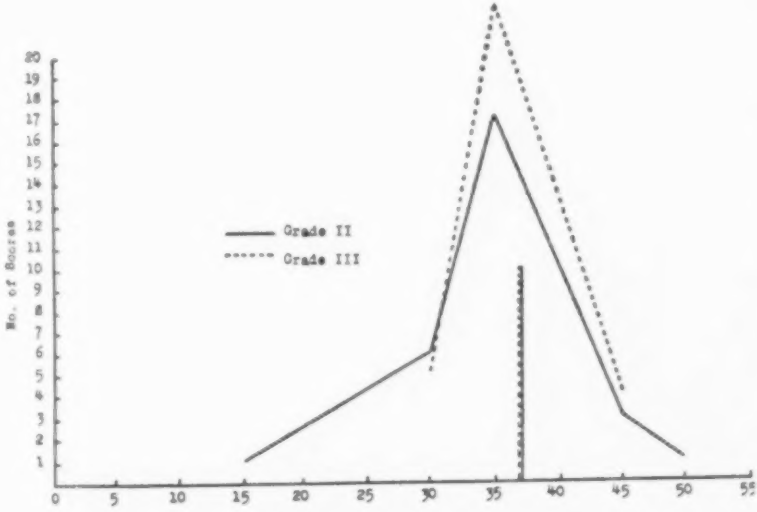

FIGURE 2. DISTRIBUTION OF SCORES ON THE DETROIT TEST

When ranked on the basis of the coefficient of variation, the Dearborn and Haggerty tests change the positions which they obtained by ranking on the basis of percent of overlapping. The four remaining tests hold the same rank by both methods.

On the Pressey test the second-grade class made a higher median score than did the third-grade group. This may have

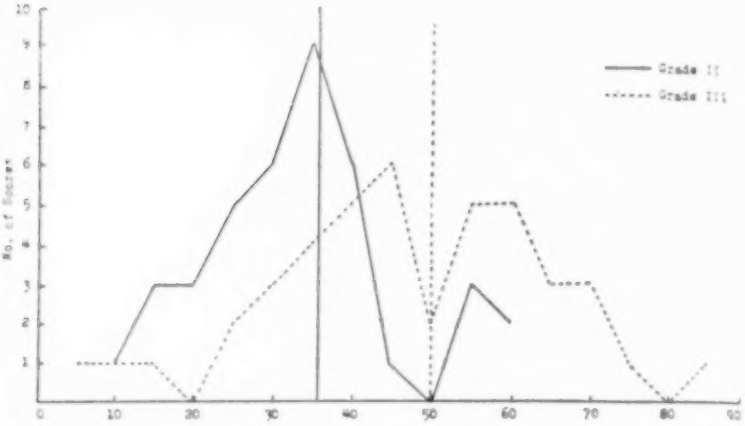

FIGURE 3. DISTRIBUTION OF SCORES ON THE HAGGERTY TEST

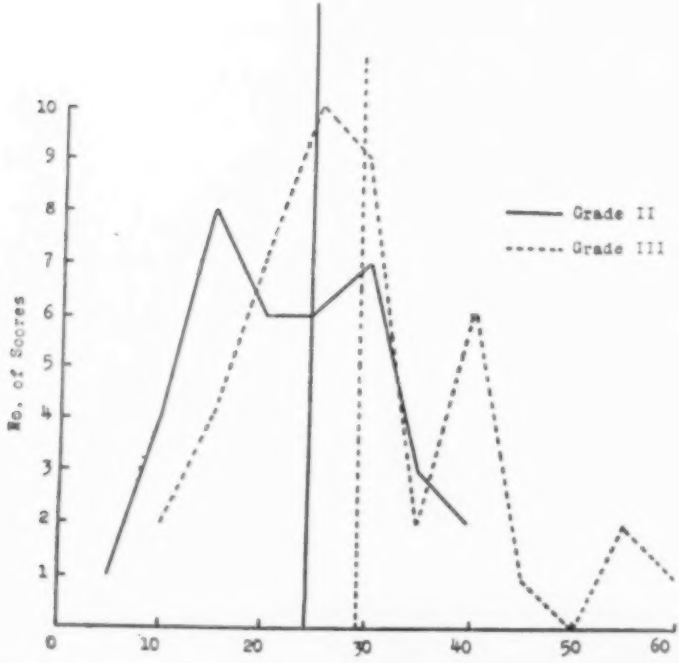

FIGURE 4. DISTRIBUTION OF SCORES ON THE MYERS TEST

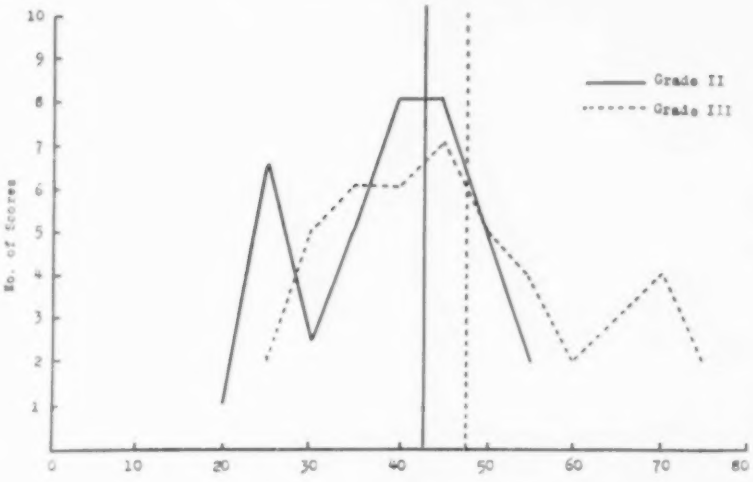

FIGURE 5. DISTRIBUTION OF SCORES ON THE OTIS TEST

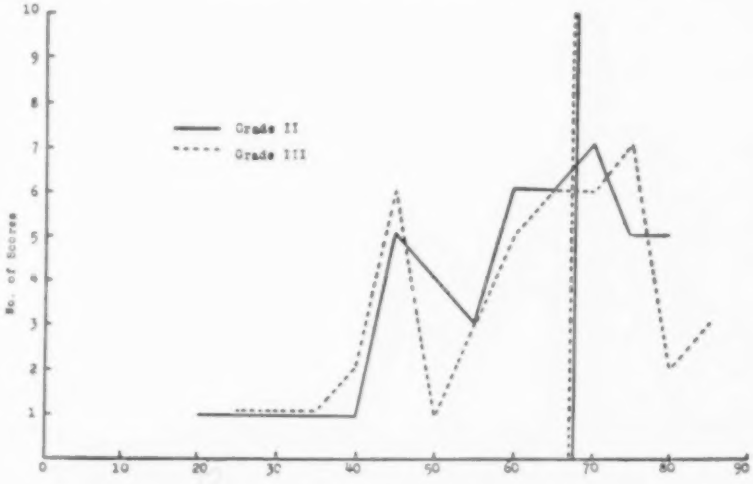

FIGURE 6. DISTRIBUTION OF SCORES ON THE PRESSEY TEST

been due to a possibly undetected error in administering the test. The standard scores given by the author, however, show a difference between second- and third-grade norms of 8 as compared with 17 between first- and second-grade norms. The De-

troit test was never intended for use in the second and third grades and obviously could not be used for classification purposes. The overlapping of the Myers and Otis tests is too large for satisfactory use in classification. The coefficients of variation for the Haggerty and Dearborn tests are larger than Henmon and Streitz obtained from testing grades I and II.

Administrative Features

The administrative features of a test are an important factor influencing the superintendent's decision as to the test to use. There are at least five administrative features that may be determined with a reasonable degree of accuracy. These are: cost per pupil, time required to give the test, time required to score the test papers and do the clerical work involved, the ease of administering, and the objectivity of scoring. All other matters being considered, the cost per pupil will most often determine the superintendent's choice of tests. The amount of time required for administering the test is important from the standpoint of the time required of examiner and children. The clerical work involved in scoring the test papers and tabulating the data obtained necessarily limits the testing program. By ease of administering and objectivity of scoring is meant in reality that the test is "fool proof," because other things being equal a test should be selected which can be administered and scored with the least possible error.

The ranking given in Table V was based on data obtained as follows: (1) the cost per pupil taken from the latest catalogues and price lists issued by publishers; (2) the total time required for administering the test, including the giving of directions and pupils' working period based upon the author's claims and experience in giving the several tests; (3) the time for scoring test papers based on the average of the estimates of the three examiners who scored the papers; (4) the ease of administering based upon the average of ranking given by two examiners responsible for administering the test; (5) the ease of scoring based upon the average of ranks given by three examiners.

The first consideration is entirely objective, the second is less objective, because different examiners will vary somewhat in their rate of reading or giving directions or distributing papers. The estimate of time required to score the test should be based, not upon that required by the author or the clerical assistant who works continuously for weeks on the same test, but on the time that will be required by teachers and clerical workers who give only a part of their time to scoring test papers and who will probably work only a few days on any one test during the school year. The ranking for objectivity of scoring and ease of administering the test is purely subjective and in this particular case is subject to considerable error. Both factors, however, must be considered in evaluating any test.

TABLE V. RANKING OF THE SIX TESTS ACCORDING TO ADMINISTRATIVE FEATURES

Test	Cost	Time of Giving	Time of Scoring	Ease of Administering	Objectivity of Scoring	Composite Ranks
Dearborn	2	6	6	6	6	6
Detroit	4.5	3	2	1	3	2
Haggerty	6	4	5	4.5	5	5
Myers	3	1.5	1	4.5	2	3
Otis	4.5	5	4	3	4	4
Pressey	1	1.5	3	2	1	1

By giving each of the five features equal weight and averaging the five ranks we obtain a composite ranking on administrative features of the several tests as follows: Pressey, one; Detroit, two; Myers, three; Otis, four; Haggerty, five; and Dearborn, six. This method of rating tests could be materially improved by having a large number of competent workers, who have had considerable experience with group intelligence tests, rate the several factors which include any subjective element.

Summary

Although the Detroit test was not originally intended for use in grades II and III, it ranked third in correlation with the composite scores and second in correlation with the Binet test and

administrative features. It was fifth in discriminative capacity. Apparently the test would be quite serviceable for a classroom teacher to use in rating the pupils within her class, but it is not serviceable as a means of classifying pupils according to grade. The Pressey test ranked sixth in the composite scoring, fifth with the Binet, sixth in discriminative capacity, and first in administrative features. Only when the administrative features are the most important factor should the Pressey test be considered. The Dearborn test ranks second in its correlation with the composite score, fourth with Binet, first in discriminative capacity, and sixth in administrative features. The Haggerty ranks fourth with composite scores, third with Binet, second in discriminative capacity, and fifth in administrative features. The Myers ranks fifth in composite scores, sixth with Binet Scale, third in discriminative capacity, and third in administrative features. The Otis ranks respectively one, one, four, and four. This article does not attempt to determine the weight that should be given on each ranking. It appears, however, there is little choice between the Otis, Haggerty, and Dearborn tests for grades II and III. The Myers and Pressey tests serve, if not a less important, at least a different purpose.

In conclusion, teachers and supervisory officers are looking to research workers for advice as to which tests to use. Advice is rendered either on the basis of opinion or established fact. There is need for agreement among research workers as to what criteria should be used for establishing the choice of tests and then for a series of investigations under the direction of unbiased leadership to determine the facts obtained from the application of the criteria.

PROPOSALS FOR STANDARIZING MEASUREMENT IN EDUCATION: II. THE COEFFICIENT OF APPLICATION

HARRY S. WILL
Toledo, Ohio

When kental scores are computed from standard norms of dispersion and central tendency, the direct comparison of the educational status of pupils in different classes, grades, or communities is made easy. If a pupil's rating in grade V is 68*k*, and in grade VI it becomes 78*k*, we know that measurable advancement has been made in the standing of the pupil relative to his group. Or if a pupil in a given community is rated *xk*, and a second pupil in another community is also rated *xk*, we know that the status of the two in the generalized educational community is identical. Chapman's notation may be employed to distinguish between the ratings of pupils in different grade or age groups. Thus, denoting grade by a Roman numeral and age by an Arabic, a designation of VIII-70*k* would indicate a score based on eighth-grade norms, while 8-70*k* would mean that the kental score was in terms of eighth-year norms.

With regard to the measurement of groups, it is further evident that results of tests in different subjects and results of different tests in the same subject are rendered more comparable when averages are stated in terms of kentals, provided again that transversions are based on standard norms. For example, if the median of community A in arithmetic problems is stated as 46*k*, and the median in arithmetic fundamentals is stated as 53*k*, the degree of variation from the norm in the two cases is more easily compared than it would be from a mere statement of original medians. And likewise, if community B showed a median of 51*k* in reading comprehension and a median of 48*k* in reading rate, it would at once be manifest that community B varied less from the norm with respect to reading abilities than did community A with respect to arithmetic abilities. Such comparisons could hardly be made from original medians.

Turning now to the matter of measuring the extent to which a pupil applies his intellectual abilities to the performance of an assigned educational task, let us review the mode of computing the coefficient designed to serve this end. We begin with noting that were the correlation between educational attainment and intelligence perfect, correspondence between a pupil's kental attainment score and his kental intelligence score would be exact. That is, an intelligence of xk would invariably yield an attainment of xk. If, therefore, we find that an intelligence of xk yields an attainment of $(x \pm u)k$, we are warranted in assuming that factors other than intelligence have operated to produce the difference of $\pm u$. Denoting the coefficient of application by a, educational attainment stated in kentals by ke, intellectual ability stated in kentals by ki, and perfect correspondence between attainment and intelligence by 100, we measure the operation of the non-intelligence factors by the equation

$$a_1 = 100 + [(ke) - (ki)]. \qquad (1)$$

Thus, if we substitute the values $(ke) = 85$ and $(ki) = 75$ in the formula, a is obviously equal to 110. If these values for ke and ki are interchanged, a will equal 90. Theoretically, since both ke and ki range between the values 0 and 100, a may range anywhere between the values 0 and 200, though practically this range will be much less.

The term "coefficient of application" is used here instead of the term "coefficient of learning" employed in the first paper[1] for the reason that it seems more aptly descriptive of the function of the measure. While the coefficient of application performs the same function as the achievement quotient, it is to be noted that the achievement quotient involves the computation of the achievement age and the mental age, while the coefficient of application involves neither.

An important criticism which may be urged against the coefficient of application is concerned with its precision. The fundamental assumption on which the computation of the co-

[1] Will, Harry S. "A Method of Commensurating Mental Measurements." *Journal of Educational Research.* 5:139-152. February, 1922.

efficient, a_1, is based, is that correlation between intelligence and educational attainment is generally perfect, the thing measured being *deviation from perfect connection*. In other words, the coefficient, a_1, measures the dislocation of a given educational score from perfect connection with its associated intelligence score, due to the operation of factors other than intelligence. Now, as a matter of fact, educational measurements are by no means perfectly correlated with intelligence, reading coming highest with a correlation of around 0.6. In reality, then, what the coefficient measures is, not deviation from *perfect* connection, but deviation from *known* connection. Consequently, in estimating an educational score from its intelligence correlate by means of the regression equation, a certain error is involved which is neglected by equation (1). This error is measured by the relation $s_x = \sigma_x \sqrt{1 - r^2}$. This expression, which is termed the standard or root-mean-square error of regression, may be regarded as defining the approximate limit of accuracy with which an educational score may be inferred from its associated intelligence score. Therefore, since the given educational score, viewed as a correlate of an intelligence score, is not (ke), but $(ke) \pm s$, the formula for computing the coefficient of application should read

$$a_2 = 100 + [(ke) \pm s] - (ki) = a_1 \pm s. \qquad (2)$$

In the equation of s_x it will be noted that the values of both σ and r are to be computed from the kental transversions and not from the original scores. This does not materially affect the value of r, but it does affect the value of σ, as will now be shown.

Let σ_0 denote the standard deviation of the original distribution; σ_1, the kental transversion of σ_0; and σ_2, the standard deviation of the kental transversions of the original scores. From the equation of the standard range,

$$R_1 = 6.0705\ \sigma_0 = 100k = R_2,$$

we get

$$\sigma_0 = \frac{R_2}{6.0705} = \frac{100k}{6.0705} = 16.5k = \sigma_1$$

Thus we see that, when expressed in kentals, the standard deviation, σ_0, of the original distribution is always equal to $16.5k$. Further, from the equation,

$$k = 50 + \frac{100d}{R_1}$$

we get

$$d = \frac{R_1 (k - 50)}{100}$$

Whence it is evident that

$$\sqrt{\frac{\Sigma d^2}{n}} = \sigma_0 = \frac{R_1}{100} \sqrt{\frac{\Sigma (k - 50)^2}{n}}$$

But, noting that 50 is set equal to the mean of the distribution and that $k - 50$ gives the deviation from the mean, we see that the expression $\sqrt{\frac{\Sigma (k - 50)^2}{n}}$ is the standard deviation, σ_2, of the kental transversions of the original values. It follows, then, since $R_1 = 6.0705\ \sigma_0$, that $\sigma_2 = \frac{100\ \sigma_0}{6.0705\ \sigma_0} = 16.5$. The value, then, to be substituted for σ in the correction formula $s_c = \sigma_x \sqrt{1 - r^2}$ is always 16.5, no matter what the standard deviation of the original distribution may be.

Where the most exact results are required, a further correction of the value of σ must be made, for it will be noted that both σ and r in the equation of the regression error are subject to the fluctuations of sampling. According to the theory of probabilities, the true value of σ lies somewhere within the limits given by the expression $\pm \frac{3\sigma}{\sqrt{2n}}$, and the true value of r lies within the limits $\pm \frac{3 (1 - r^2)}{\sqrt{n}}$. Let us designate the regression error s_x, when corrected for the fluctuations of sampling, by the symbol e, the equation of which is given by making the appropriate substitutions in the equation of s_x. We have, therefore,

$$e = \left[\sigma \pm \frac{3\sigma}{\sqrt{2n}} \right] \left[1 - \left\{ r \pm \frac{3 (1 - r^2)}{\sqrt{n}} \right\}^2 \right]^{\frac{1}{2}}$$

In this equation we note that the function e varies directly with σ and inversely with r. The significant limits of e, therefore,

are given when σ takes its maximum value and r its minimum value. We may, then, write the function c as follows:

$$c = \left[\sigma + \frac{3\sigma}{\sqrt{2n}} \right] \left[1 - (r\sqrt{n} - 3 + 3r^2)^2 . \frac{1}{n} \right]^{\frac{1}{2}}$$

We now write the equation of the coefficient with the second correction as

$$a_3 = 100 + [(kc) \pm c] - (ki) = a_1 \pm c. \qquad (3)$$

When, in the equation of s or c, $\sigma = 16.5$ and $r = 0.6$, $s = 13.2$. If, when employing the correction c, n takes the values $n = 50$, $n = 500$, and $n = 1000$, then $c = 20.1$, 15.3, and 14.8 for the successive values of n. We have, then, when $(kc) = 70$ and $(ki) = 80$: $a_1 = 100 + 70 - 80 = 90$; $a_2 = a_1 \pm s = 90 \pm 13.2$; $a_3 = a_1 \pm c = 90 \pm 20.1$, $(n = 50)$; $a_3 = a_1 \pm c = 90 \pm 15.3$, $(n = 500)$; $a_3 = a_1 \pm c = 90 + 14.8$, $(n = 1000)$. Since a (kc) score can be deduced from a given (ki) score only within the limits of the regression errors s and c, according to the degree of approximation we desire to attain, no inference can be drawn from values of a falling within these limits as regards the efficiency of intelligence in producing the given educational score.

When it is desired to use the coefficient as a general check on the effectiveness of instruction, the average score of the instructional unit may replace the score of the individual pupil as the basis of comparison between intelligence and educational attainment. In this case, the value of r to be used in the computation of the regression error is to be computed from the group averages and not from individual scores. Where accuracy is of moment, each squared deviation figuring in the computation of the standard deviation and each xy product figuring in the product sum is to be weighted for the number of individuals in the instructional group yielding the given average measurement.

The argument may be brought forward that measures such as the coefficient of application or the achievement quotient are spurious innovations into the field of educational measurement. After all, is it worth while to measure such a quality as application? Is it not a form of over-learning to require of a bright

pupil that he put a certain percent of his ability into the doing of an assigned task? Is it not enough that he measure up to the mark set for advancement into the next higher instructional unit? And again, is it desirable that the educational system absorb the pupil's energies to the extent that natural initiative is given no chance to flourish? These queries are pertinent. The quest for novel formulas is altogether idle if it leads to no practical result.

There are, however, rejoinders which may be made to the questions stated. We may hold it valuable information to learn that inattentiveness or mischievousness on the part of a pupil are explainable on the ground of inadequate stimulation. If we are interested merely in the attainment of minimum standards, the coefficient might still inform us when an increase in the variety of work would be profitable. In this way the pupil's power of initiative might be developed rather than dwarfed. Moreover, as has been pointed out editorially in this journal, acceleration is properly to be regarded as definitely associated with enrichment through the enlarged opportunities which it brings. At any rate, if the task set for the pupil be too light and there develop an attitude and habit of positive or relative indolence, it is properly the function of professional education to correct the condition. The situation of the pupil who, with poor capacities, is in reality over-reaching himself is better understood when his coefficient is known. And, lastly, the measure of application ought to be of real service to the judicious administrator in his estimation of instructional efficiency. A workman may do poor work with good materials or good work with poor.

A SERVICEABLE $P.E._r$ TABLE

HERBERT A. TOOPS
AND
ZAIDA F. MINER
Teachers College, Columbia University

In the formula for the probable error of the Pearson correlation coefficient,

$$P.E._r = .6745 \frac{1 - r^2}{\sqrt{N}}, \qquad (1)$$

r may take any value from -1.00 to $+1.00$ inclusive; N may take any value from 2 to ∞; and consequently $P.E._r$ may range between zero and .48, (i.e., 0 and $.6745/\sqrt{2}$.) Due to the extremely large range of N, it is impracticable to construct a table which will yield from r and N as coordinates the $P.E._r$ for even the commonly used combinations of r and N. Whenever such tables are used, it becomes necessary to interpolate N for almost every value of $P.E._r$ one obtains.[1] Such interpolation makes for errors and uncertainty.

The table below requires no interpolation unless one desires $P.E._r$'s correct to more than the nearest second decimal place, or to take r's to finer units than .01. This desirable feature is obtained by manipulating the formula for $P.E._r$ so that we solve for the largest value of N which (for a given r) will yield a given probable error. Thus for a given correlation coefficient of exactly .10, 21 is the largest N which, to the nearest .01, yields a $P.E._r$ of .15; and 24 is the largest value of N which yields a $P.E._r$ of .14 when taken to the nearest .01. So, when $r = .10$, all values of N between 22 and 24 inclusive yield $P.E._r$'s between the limits .1449 and .1351, all of which are called .14.

By an inspection of the values of $P.E._r$ for $r = 0$ (which of all correlation coefficients has the maximum $P.E._r$), it will be found that $P.E._r$'s of over .20 occur for only N's of ten cases or

[1] Rugg, H. O. *Statistical Methods Applied to Education*, New York, Houghton Mifflin Co., 1917, Table X, p. 404.

less.[2] Consequently, if for each given value of r (in .01's) we compute the *maximum* value of N which will just escape association with a next lower $P.E._r$ for the range in $P.E._r$ of .01 to .20 inclusive, we shall be in a position to build an extremely useful table for determining the $P.E._r$. Solving formula (1) for N,

$$N = \left(\frac{.6745}{P.E._r}\right)^2 .(1-r^2)^2 = \frac{.45495025}{P.E._r^2} .(1-r^2)^2. \qquad (2)$$

An examination of this formula shows that N gets larger as r gets smaller. Consequently in the table below the N's decrease from the top of the table ($r = 0$) towards the bottom. Furthermore, N will be a maximum when $P.E._r$ is a minimum; consequently, to determine the *largest* possible value of N (in integers) which will yield a given $P.E._r$ of, say, .15, we must take $P.E._r$ of the *lowest* value which, to the nearest .01, can be called .15; i.e., since all $P.E._r$'s of .14500001 to .15499999 inclusive will be called .15 when we throw away the decimal places beyond the second place, we must take $P.E._r$ of the value of .145 in Formula (2). The published table then results. An r-ordinate of .49 is .490000.

To Use the Table in Finding $P.E._r$

Example 1.—Required the $P.E._r$ when $r = .20$ and $N = 76$. In the row for $r = .20$ trace across to the right until the *first* value of N *larger than* 76 is located. This entry is 99. Immediately above the 99 entry at the top of the page is the $P.E._r$ sought, namely .07. The entry 99 just located in reality thus stands for the range of values, 75-99 inclusive, all of which have the same $P.E._r$ of .07, to the nearest second decimal place. When $r = .20$, if N becomes larger than 99, $P.E._r$ drops to .06 or less; if N becomes 74 or less, $P.E._r$ becomes .08 or more.

Example 2.—Required the $P.E._r$ when $r = .14$ and $N = 32$. In similar manner the entry 33 in row .14 guides us to $P.E._r = .12$. For $r = .14$, if N becomes larger than 33, $P.E._r$ drops to .11 or less; while if N is 27 or less, $P.E._r$ is .13 or more.

[2] When N is small the ordinary formula does not hold, and the denominator of formula 1 must be decreased in order to obtain the $P.E._r$.

$P.E._r$ TABLE

$\pm r$	Largest Value of N which has the $P.E._r$:—																				$\pm r$
	.20	.19	.18	.17	.16	.15	.14	.13	.12	.11	.10	.09	.08	.07	.06	.05	.04	.03	.02	.01*	
.00	11	13	14	16	18	21	24	29	34	41	50	62	80	107	150	224	371	727	2022	18198	.00
.01	11	13	14	16	18	21	24	29	34	41	50	62	80	107	150	224	371	727	2021	18194	.01
.02	11	13	14	16	18	21	24	29	34	41	50	62	80	107	150	224	371	727	2020	18183	.02
.03	11	13	14	16	18	21	24	29	34	41	50	62	80	107	150	224	370	726	2018	18165	.03
.04	11	13	14	16	18	21	24	29	34	41	50	62	80	107	149	223	370	725	2015	18139	.04
.05	11	13	14	16	18	21	24	28	34	41	50	62	80	107	149	223	369	724	2011	18107	**.05**
.06	11	13	14	16	18	21	24	28	34	40	50	62	80	107	149	223	368	722	2007	18067	.06
.07	11	13	14	16	18	21	24	28	34	40	49	62	80	106	148	222	367	720	2002	18020	.07
.08	11	13	14	16	18	21	24	28	33	40	49	62	79	106	148	221	366	718	1996	17965	.08
.09	11	13	14	16	18	21	24	28	33	40	49	61	79	105	147	221	365	716	1988	17900	.09
.10	11	13	14	16	18	21	24	28	33	40	49	61	79	105	147	220	363	713	1981	17835	**.10**
.11	11	12	14	16	18	21	24	28	33	40	49	61	78	105	146	219	362	710	1973	17760	.11
.12	11	12	14	16	18	21	24	28	33	40	48	61	78	104	146	218	360	707	1964	17677	.12
.13	11	12	14	16	18	20	24	28	33	39	48	60	78	104	145	217	358	703	1954	17588	.13
.14	11	12	14	16	18	20	23	27	33	39	48	60	77	103	144	215	356	699	1943	17491	.14
.15	11	12	14	15	18	20	23	27	32	39	48	60	77	102	143	214	354	695	1932	17388	**.15**
.16	11	12	14	15	17	20	23	27	32	39	47	59	76	102	142	213	352	691	1919	17278	.16
.17	11	12	14	15	17	20	23	27	32	38	47	59	76	101	141	211	350	686	1906	17161	.17
.18	11	12	13	15	17	20	23	27	32	38	47	58	75	100	140	210	347	681	1893	17037	.18
.19	11	12	13	15	17	20	23	27	31	38	46	58	75	100	139	208	345	676	1878	16907	.19
.20	11	12	13	15	17	19	23	26	31	38	46	58	74	99	138	207	342	670	1863	16771	**.20**
.21	10	12	13	15	17	19	22	26	31	37	46	57	73	98	137	205	339	665	1847	16628	.21
.22	10	12	13	15	17	19	22	26	31	37	45	57	73	97	136	203	336	659	1831	16479	.22
.23	10	11	13	14	16	19	22	26	30	37	45	56	72	96	134	201	333	652	1813	16323	.23
.24	10	11	13	14	16	19	22	25	30	36	44	55	71	95	133	199	329	646	1795	16161	.24
.25	10	11	13	14	16	19	21	25	30	36	44	55	71	94	132	197	326	639	1777	15994	**.25**
.26	10	11	12	14	16	18	21	25	29	35	43	54	70	93	130	195	322	632	1757	15820	.26
.27	10	11	12	14	16	18	21	25	29	35	43	54	69	92	129	193	319	625	1737	15641	.27
.28	10	11	12	14	16	18	21	24	29	35	42	53	68	91	127	190	315	618	1717	15456	.28
.29	10	11	12	14	15	18	20	24	28	34	42	52	67	90	126	188	311	610	1696	15265	.29
.30	9	11	12	13	15	17	20	24	28	34	41	52	66	89	124	186	307	602	1674	15069	**.30**
.31	9	10	12	13	15	17	20	23	28	33	41	51	66	87	122	183	303	594	1652	14868	.31
.32	9	10	11	13	15	17	20	23	27	33	40	50	65	86	121	181	299	586	1629	14661	.32
.33	9	10	11	13	15	17	19	23	27	32	40	50	64	85	119	178	294	578	1605	14450	.33
.34	9	10	11	13	14	16	19	22	26	32	39	49	63	84	117	175	290	569	1581	14233	.34
.35	9	10	11	12	14	16	19	12	26	31	38	48	62	82	115	172	285	560	1556	14012	**.35**
$\pm r$	**.20**	**.19**	**.18**	**.17**	**.16**	**.15**	**.14**	**.13**	**.12**	**.11**	**.10**	**.09**	**.08**	**.07**	**.06**	**.05**	**.04**	**.03**	**.02**	**.01**	$\pm r$

*$P.E._r$ for all N's larger than indicated in this column is zero.

$P.E._r$ TABLE (*Continued*)

$\pm r$	Largest Value of N which has the $P.E._r$:—																$\pm r$
	.20	.19	.18	.17	.16	.15	.14	.13	.12	.11	.10	.09	.08	.07	.06	.05	.04	.03	.02	.01**	
.36	9	10	11	12	14	16	18	22	26	31	38	47	61	81	113	170	281	551	1531	13786	.36
.37	8*	9	11	12	14	16	18	21	25	30	37	46	60	80	112	167	276	542	1506	13556	.37
.38	8	9	10	12	13	15	18	21	25	30	36	46	59	78	110	164	271	532	1480	13321	.38
.39	8	9	10	12	13	15	17	20	24	29	36	45	58	77	108	161	267	523	1453	13083	.39
.40	8	9	10	11	13	15	17	20	24	29	35	44	57	75	105	158	262	513	1426	12840	**.40**
.41	8	9	10	11	13	14	17	20	23	28	34	43	55	74	104	155	257	503	1399	12594	.41
.42	8	9	10	11	12	14	16	19	23	27	34	42	54	73	102	152	251	493	1371	12344	.42
.43	7*	8	9	11	12	14	16	19	22	27	33	41	53	71	99	149	246	483	1343	12090	.43
.44	7*	8	9	10	12	14	16	18	22	26	32	40	52	70	97	146	241	473	1314	11833	.44
.45	7	8	9	10	12	13	15	18	21	26	32	40	51	68	95	142	236	462	1286	11574	**.45**
.46	7	8	9	10	11	13	15	18	21	25	31	39	50	66	93	139	230	452	1256	11311	.46
.47	7	8	9	10	11	13	15	17	20	25	30	38	49	65	91	136	225	441	1227	11046	.47
.48	7	7	8	9	11	12	14	17	20	24	29	37	47	63	89	133	219	431	1197	10778	.48
.49	6*	7	8	9	10	12	14	16	19	23	29	36	46	62	86	129	214	420	1167	10508	.49
.50	6*	7	8	9	10	12	14	16	19	23	28	35	45	60	84	126	208	409	1137	10236	**.50**
.51	6	7	8	9	10	11	13	15	18	22	27	34	44	58	82	122	203	398	1106	9962	.51
.52	6	7	7	8	10	11	13	15	18	21	26	33	43	57	80	119	197	387	1076	9687	.52
.53	6	6	7	8	9	11	12	15	17	21	26	32	41	55	77	116	192	376	1045	9410	.53
.54	6	6	7	8	9	10	12	14	17	20	25	31	40	54	75	121	186	365	1014	9132	.54
.55	5*	6	7	8	9	10	12	14	16	20	24	30	39	52	73	109	180	354	983	8853	**.55**
.56	5*	6	7	7	8	10	11	13	16	19	23	29	38	50	70	105	174	342	952	8573	.56
.57	5	6	6	7	8	9	11	13	15	18	22	28	36	49	68	102	169	331	921	8293	.57
.58	5	5	6	7	8	9	10	12	15	18	22	27	35	47	66	98	163	320	890	8013	.58
.59	5	5	6	7	8	9	10	12	14	17	21	26	34	45	63	95	157	309	859	7733	.59
.60		5	6	6	7	8	10	11	14	16	20	25	33	44	61	92	152	298	828	7453	**.60**
.61		5	5	6	7	8	9	11	13	16	19	24	31	42	59	88	146	286	797	7174	.61
.62		5	5	6	7	8	9	11	13	15	19	23	30	40	56	85	140	275	766	6896	.62
.63			5	6	6	7	9	10	12	15	18	22	29	39	54	81	135	264	735	6619	.63
.64			5	5	6	7	8	10	11	14	17	21	28	37	52	78	129	253	704	6343	.64
.65				5	6	7	8	9	11	13	16	21	26	35	50	74	123	242	674	6069	**.65**
.66				5	6	6	7	9	10	13	16	20	25	34	47	71	118	231	644	5796	.66
.67				5	5	6	7	8	10	12	15	19	24	32	45	68	112	221	614	5526	.67
.68					5	6	7	8	9	11	14	18	23	31	43	64	107	210	584	5259	.68
.69					5	5	6	7	9	11	13	17	22	29	41	61	101	199	554	4994	.69
.70						5	6	7	8	10	13	16	21	28	39	58	96	189	525	4733	**.70**
$\pm r$	.20	.19	.18	.17	.16	.15	.14	.13	.12	.11	.10	.09	.08	.07	.06	.05	.04	.03	.02	.01	$\pm r$

* True $P.E._r$ is .21.

** $P.E._r$ for all N's larger than indicated in this column is zero.

True $P.E._r$ is .21.

**$P.E._r$ for all N's larger than indicated in this column is zero.

$P.E._r$ TABLE (*Continued*)

±r	Largest Value of N which has the $P.E._r$:—																	±r
	.20	.19	.18	.17	.16	.15	.14	.13	.12	.11	.10	.09	.08	.07	.06	.05	.04	.03	.02	.01*	
.71						5	6	7	8	10	12	15	19	26	36	55	91	179	497	4475	.71
.72						5	5	6	7	9	11	14	18	24	34	52	86	168	468	4220	.72
.73							5	6	7	9	10	13	17	23	32	49	81	158	441	3970	.73
.74							5	5	7	8	10	12	16	22	30	45	76	148	413	3724	.74
.75								5	6	7	9	12	15	20	28	43	71	139	387	3483	**.75**
.76								5	6	7	8	11	14	19	26	40	66	129	360	3246	.76
.77									5	6	8	10	13	17	24	37	61	120	335	3015	.77
.78									5	6	7	9	12	16	23	34	56	111	310	2790	.78
.79										5	7	8	11	15	21	31	52	102	285	2571	.79
.80										5	6	8	10	13	19	29	48	94	262	2358	**.80**
.81											5	7	9	12	17	26	43	86	239	2152	.81
.82											5	6	8	11	16	24	39	78	217	1953	.82
.83												6	7	10	14	21	35	70	195	1761	.83
.84												5	7	9	13	19	32	63	175	1577	.84
.85													6	8	11	17	28	56	155	1401	**.85**
.86													5	7	10	15	25	49	137	1233	.86
.87														6	8	13	21	43	119	1075	.87
.88														5	7	11	18	37	102	926	.88
.89															6	9	16	31	87	786	.89
.90															5	8	13	26	72	656	**.90**
.91																6	10	21	59	537	.91
.92																5	8	17	47	429	.92
.93																	6	13	36	332	.93
.94																	5	9	27	246	.94
.95																		6	19	172	**.95**
.96																			12	111	.96
.97																			7	63	.97
.98																				28	.98
.99																				7	.99
1.00																				0	**1.00**
±r	20	19	18	17	16	15	14	13	12	11	10	09	08	07	06	05	04	03	02	01	±r

*$P.E._r$ for all N's larger than indicated in this column is zero.

Example 3.—What is the least value of N which has zero $P.E._r$ (to the nearest .01) when $r = .63$? Since the *maximum* value of N which has a $P.E._r$ of .01 is 6619, and since the $P.E._r$ for all values of N from 6620 to ∞ is .00, the value of N sought is 6620. And 6620 is the smallest number of cases which, for $r = 63$, has a $P.E._r$ of .00.

The Use of the Table in Building Tables for Publication

The table may be used as a convenient means of building up specific $P.E._r$ tables to be published as adjuncts to tables of intercorrelation coefficients. The correlation coefficients in an intercorrelation table are ordinarily based on a comparatively few different N's. Hence economy of space and printing cost may be secured by printing as a footnote to such an intercorrelation table a $P.E._r$ table modeled after the following table.

Footnote to Table XXX. The probable errors of the correlation coefficients of the above table may be readily determined from the following table (negative r's have the same probable errors as positive ones of the same absolute magnitude):

Test Group	N	Range of r's having the respective $P.E._r$'s:—									
		.09	.08	.07	.06	.05	.04	.03	.02	.01	.00
A	57	.00-.22	.23-.40	.41-.52	.53-.61	.62-.70	.71-.77	.78-.84	.85-.91	.92-.97	.98-1.00
B	102	.....	.....	.00-.16	.17-.42	.43-.57	.58-.68	.69-.79	.80-.88	.89-.96	.97-1.00
C	134	.....	.....	.....	.00-.23	.24-.47	.48-.63	.64-.75	.76-.86	.87-.95	.96-1.00
D	200	.....	.....	.....	.....	.00-.23	.24-.51	.52-.68	.69-.82	.83-.94	.95-1.00
E	386	.....	.....	.....	.....	.....	.....	.00-.52	.53-.75	.76-.92	.93-1.00

The maximum probable error to include in each such table may be determined by noting from the large table the $P.E._r$ corresponding to the smallest N-group for the smallest (absolute magnitude) r which occurs in that group in the intercorrelation table. The minimum $P.E._r$ need be carried no further than to include the maximum r for each group respectively. Thus, if in Group C above the correlations in the intercorrelation table range between .23 and .78, one needs to give the r ranges corresponding to the probable errors, .06, .05, .04, .03, .02 only. These are the only $P.E._r$'s which occur when N is 134 and r has the range .23 to .78.

The use of the above footnote table is obvious from its construction. In Group B, for instance, all correlations between .43 and .57 inclusive have the same probable error, .05.

Editorials

ANNUAL AND SEMI-ANNUAL PROMOTION

Among the suggestions for a more flexible promotion system the semi-annual plan has many adherents. This arrangement contemplates a regrouping of pupils twice a year. It has superseded annual promotions in many—perhaps in a majority—of progressive school systems. The most effective argument for its adoption appears to have been the supposed reduction in retardation which it would affect.

This reduction has apparently failed to materialize. Comparisons between school systems on the basis of annual and semi-annual promotion have failed to reveal a superiority for the latter. Moreover, no school system upon passing from an annual to a semi-annual basis has, so far as we are aware, reduced its retardation under circumstances which would permit the reduction to be fairly attributed to the change of system.

If you ask the school executive how he expects this reduction to take place he will tell you that it is to be brought about by two tendencies each of which he assumes will operate to advance pupils more rapidly when promotion periods are more frequent. One of these is a greater willingness on the part of teachers to promote pupils incidentally—that is, at other than regular promotion times. The other is their greater willingness to make double promotions at the end of the semester than at the end of the year. Both tendencies are based on the reasonable belief that a pupil can advance an extra half year more readily than he can advance an extra year.

But these tendencies will not secure a net superiority in respect to rate of promotion unless they are more powerful than tendencies working in the opposite direction. We believe that they will not be more powerful than these countertendencies unless

special effort is made to bring them into full effect. In other words, in this as in other matters of school administration, it is not enough merely to set up machinery. Machinery will not run itself. If a superintendent expects to get results from a particular type of organization—whether it is a promotion scheme, a junior high school, or a work-study-play system—he must supply it with directive power, study its workings, surround it with favorable conditions, and remove the hindrances to its success. If he expects extra promotions to take place, he must give them continuous supervisory attention. Without this, teachers will not adopt in a large way a policy of selecting pupils for incidental and double promotion. As a plan to secure more rapid advancement of pupils, the semi-annual system of promotions has certain disadvantages, and in order to overcome them they must be given serious consideration.

In the first place, there is an undoubted tendency among teachers to regard failure as less serious when it involves repeating the work of half a year than they do when it involves the loss of a year in advancement. This tendency leads to the holding back of more pupils at promotion time, if the school system promotes regularly every half year. Thus the promotion rate is likely to be lowered.

In the second place, even with the same promotion rate the semi-annual plan is inferior as a means of bringing capable pupils forward. Just as interest when compounded semi-annually yields a larger return than the same rate does when compounded annually, so the promotion rate which is applied twice a year yields a larger crop of failures. Moreover, the difference is surprisingly large.

Consider, for example, the effect of an 80 percent promotion rate successively applied to one hundred pupils who may be supposed to have entered school last September (1923). If they belong in City A where annual promotions prevail, 80 of them will enter the second grade in September 1924. If, however, they belong in City B where 80 percent of the pupils are promoted both in February and June, only 64 of them will enter the

second grade in September 1924. Of course, there are compensations in favor of the semi-annual plan. Those who fail will have another chance in half a year. Two groups each year—the second one in February—will enter the second grade. In other words the comparison between the two systems is not entirely as indicated by the figures 80 and 64.

Yet it is true that no matter at what point one makes a cross section of the condition of the 100 children entering the first grade last September in Cities A and B, the number of retarded children will be greater in City B where promotion occurs twice a year. And we may remark that a cross section is precisely what the familiar age-grade table is. It shows conditions as they are at a given moment with respect to the grade and age—and perhaps progress—of the pupils. Merely on the basis of the *number* of accelerated, normal, and retarded pupils, without reference to the *amount* of acceleration or retardation the annual promotion plan makes the better showing.

Let us take another example. After eight years of instruction the situation that an undisturbed 80 percent promotion rate will produce in a semi-annual scheme is shown by the following grade placement of an original 100 children.

Grade.....	High Fifth	Sixth		Seventh		Eighth		Low Ninth	Total
		Low	High	Low	High	Low	High		
Pupils.....	3	5	12	20	25	21	11	3	100

According to the up-to-14 standard for graduation from the elementary school, these pupils are not nor have they ever been accelerated. Only three of them are normal, and *97 of them are retarded.*

Similarily, after eight years of school life where an inflexible annual promotion rate of 80 percent was in vogue, 100 pupils would be distributed as follows:

Grade.....	IV	V	VI	VII	VIII	IX	Total
Pupils.....	1	5	14	29	34	17	100

Here 17 pupils are normal as compared with three on the semi-annual basis. Correspondingly, 83 are retarded as compared with 97. Again, however, no one is or ever has been accelerated. Such a thing cannot happen with any promotion rate of less than 100 if it is rigidly applied, unless a pupil is already accelerated when he enters school.

In general, we shall find that the year basis of promotion tends to place more pupils in the higher grades in a given length of time. In the tenth year of the school life of the 100 pupils who were supposed to be promoted annually, 30 would be found in the first year of the high school and 14 in the second year. If we may define a "pupil-year" as one pupil taught one year, this, of course, gives us 44 pupil-years. The semi-annual promotion plan would only yield 36 pupil-years of high-school instruction in the tenth year of schooling.

Although annual promotion causes fewer children to be retarded, the amount of retardation per child is likely to be greater. In the first table we have given above, no pupil lags by more than 3½ years. In the second table, however, one child is five years and five children are four years behind their "normal" grade.

While, therefore, it is true that the two systems of promotion—annual and semi-annual—have, if undisturbed, precisely the same mathematical results in the long run, their practical significance and ease of application are by no means the same. The lesson we may derive from the foregoing discussion is that with a semi-annual promotion plan, special effort must be made with the brighter children to keep them moving forward. With the annual promotion plan the special effort must be with the slower pupils. The matter may be stated another way. The annual system tends to take better care of the bright child, while the semi-annual system favors the dull child. Neither one goes very far, however, in either direction; and the superintendent who wishes to get results will be little better off merely for a change of interval between regular promotions.

B. R. B.

Reviews and Abstracts

HANSEN, ALLEN O. *Early Educational Leadership in the Ohio Valley.* Bloomington, Illinois. Public School Publishing Company, 1923. 120 pp. (Journal of Educational Research Monographs, No. 5.)

In this little volume Mr. Hansen describes the efforts towards educational reconstruction on the part of the Western Literary Institute and College of Professional Teachers between 1829 and 1841. Chapters are devoted to the origin of the College, the program for the nationalization of schools, the movement towards state support and control, educational objectives, proposed changes in curriculum and methods, and the distribution and influence of the College of Teachers.

Mr. Hansen finds the spirit actuating the College of Teachers to be that of the New West—free, creative, individualistic, critical. Theirs is a pioneer's problem: to provide "education that would function in the development of the resources of the West, and the integration of the life of the diverse population by the American ideal." More specifically, they sought to create the proper public attitude towards education, elevate the teaching profession, promote professional acquaintance and friendly associations, collect educational data, and arrange and digest it so as to lay the foundation for a true science of education.

In accord with this program we find the *Transactions* dealing with a nationalization program, some members urging not "100-percent Americanism" but a "purely American" core of studies—this because "the most disastrous consequences would result from the ignorance and misrule of a foreign immigration." State controlled systems were agreed to be necessary, and petitions were sent to legislative bodies urging the establishment of universal education and training institutions to provide proper teachers. On the necessary changes in curriculum there were many debates, the following issues being involved:

Old versus new; classical versus scientific; ancient languages versus modern; utilitarian versus cultural education; individual versus social; indigenous versus exotic; higher versus lower; private versus public; short and superficial courses versus long and profound; simplification or reduction versus enrichment; life versus information; interests and capacities of children versus adult imposition; female versus male.

Many modern views are expressed. Likewise, in the discussion of methods such modern ideas are found: "We will here introduce the motive that we think best adapted to the point in question. This is the motive of interest;" "Imagine, for a moment, an institution where exercises like these are daily practised. Behold a teacher surrounded, not by mere mechanical beings set

in motion by the springs of memory, but by a class of *listening, thinking, speaking* pupils;" and "Nothing should be told to the child directly . . . for whatever he discovers himself, or thinks he discovers, is remembered with delight."

The study is based on the *Transactions of the College of Teachers* almost entirely, and for the most part these records are allowed to tell the story in their own straightforward way. On most questions with which the author is concerned, the *Transactions* are obviously the best, or only, authority, and the inclusion of much of this source material in its original form constitutes one of the excellences of the book; but to decide the probable influence of the organization (in the last chapter) a wider range of authorities might well have been used. No bibliography is given, but Turner's *Rise of the New West*, *The Western Academician*, Cist's *Cincinnati in 1841*, and others are referred to. The usefulness of the book for students of educational history would be greatly enhanced by an index.

Notwithstanding the above limitations, the study is a valuable contribution to the monographic literature on the educational history of the United States, especially since many students lack access to the originals from which it is drawn.

The editor and publishers of the *Educational Research Monographs* are to be complimented on the inclusion of historical studies in their series; the pecuniary reward will, doubtless, be small, but the service great, since it will hasten the day when a genuine history of education in the United States can be written.

THOMAS WOODY

University of Pennsylvania

GREGORY, CHESTER A. *Fundamentals of Educational Measurement with the Elements of Statistical Method.* New York, D. Appleton and Company, 1922. 382 pp.

"To bridge, in part, a gap between theory and practice in educational tests and measurements" is given by the author as the purpose of this book. This would lead one to assume that the theoretical basis of educational measurements would be presented in such a way that school teachers and students of education might come to see the value of such measures in their daily work. This is attempted not by telling teachers how to give and score tests, nor by showing what results have been attained through the use of tests, nor by furnishing definite criteria for the selection of specific tests for specific purposes, nor by clearly telling or illustrating how the results of a testing program may be used for purposes of classification or remedial work. Rather it is done by recounting the need for more exact measurement in education, by sketchy descriptions of the mental testing movement and its products, by explanations of some of the methods and products of achievement-test making, by briefly pointing out and illustrating measurements in other fields, and by explaining through samples what is meant by some of the terms and processes used in the statistical treatment of educational measurements.

The book covers too wide a scope to be concrete enough to stimulate teachers to attempt a testing program or to arouse in them a curiosity to find out more about tests and their values. At the same time it is not sufficiently exhaustive in its treatment of any phase of the testing movement to give the reader a thorough comprehension of it.

M. J. VAN WAGENEN

University of Minnesota

WOOLLEY, HELEN THOMPSON, AND FERRIS, ELIZABETH. *Diagnosis and Treatment of Young School Failures.* Washington, D. C., Government Printing Office, 1923. 114 pp. (Department of the Interior, Bureau of Education, Bulletin 1923, No. 1.)

Cincinnati is steadily building up a reputation for cooperativeness in community welfare work, and has shown a readiness to approach scientifically the solution of its social problems. The Vocational Bureau, the Council of Social Agencies, the recent Mental Hygiene Survey, the Public Health Federation, and, we may venture to add, the Social Unit experiment, are indications of community intelligence in Cincinnati. The present bulletin, though it bears the imprint of the Government Printing Office, was prepared under the auspices of the Helen S. Trounstine Foundation, Cincinnati, and constitutes Number 8 in the foundation's series of social studies.

The main body of this bulletin consists of sixteen readable case reports of children ranging from seven to ten and a half years of age, distributed in the four following groups: nine neglected children, four high-grade defectives, two children with special defects, and one psychopathic child.

The reports are readable because, although they detail the ratings of the children on several standard psychological tests, they also embody personal, social, and educational data which make the cases human and instructive. These children spent a long period in an observation class where they learned to like school, and received the benefit of Miss Ferris' sympathetic observation and instruction. The scope of the case studies includes home history, physical health, inheritance, and personality make-up.

The investigation involved successive psycho-metric examinations. These showed a strikingly uniform increase in intelligence quotients after the children had attended the observation class, and an equal decrease in intelligence quotients when re-examined after a year or a year and a half in a regular class.

The bulletin serves a valuable purpose by displaying the limitations of an uninterrupted intelligence quotient rating and by emphasizing the importance of personality factors even in young children. By implication the significance of the pre-school period of development and of pre-school educational adjustment is also recognized.

The section, "Discussion of Results," contains this sentence which has far-reaching import for both teachers and social workers: "Our experience with the observation class convinces us that if the task can be properly adjusted to the degree of ability of the child, it is possible for many of the inferior children to experience the same love of learning and the same desire for inde-

pendent work which is characteristic of the higher levels of accomplishment."

Although this bulletin can be read with profit by social workers, it should be peculiarly useful to primary teachers and to teachers in training. Since it is an inexpensive government bulletin this pamphlet may well have a wide circulation in normal schools. Two dollars will put twenty copies on the reserve shelf, a dime makes you a possessor.

ARNOLD GESELL

Yale University

COOPER, RICHARD W., AND COOPER, HERMANN. *Negro School Attendance in Delaware.* Newark, Delaware, University of Delaware Press, 1923. 389 pp.

Negro School Attendance in Delaware by Cooper and Cooper is a book of facts. Its appearance is attractive, its method is scientific, and it abounds in graphs and tables which tell at a glance the school status of this class of children. The problem of attendance is attacked in such a way that business men, educators, or the most humble patrons may see the working program of education in its strength and weakness. Poor attendance is a weakness, and it is the most important cause for retardation, non-promotion, high costs, and misunderstandings. Remedies designed to bring about better attendance are important; they appeal to the taxpaying public and are a challenge to the parents and others directly interested in education.

This publication comes at an opportune time, for though local in character, it will provide suggestive information for many sections of the country. It gives the facts demanded by an interested public.

JOHN W. DAVIS

Institute, West Virginia

CRAMPTON, C. WARD. *The Pedagogy of Physical Training.* New York, The Macmillan Company, 1922. 257 pp.

This book comes as the result of Doctor Crampton's long experience in teacher training and in the supervision of physical education in the schools. The material is presented in simple concise form, and should appeal to thousands of teachers in the public schools who are called upon to conduct activities in physical education without previous technical training.

In the first chapter, discussing physical training as a profession, the author points out the constantly increasing demand for trained teachers in this field, and the unusual opportunity for service which is presented. The chapter on professional qualifications is excellent and will be helpful to the teacher in any branch of education.

In the opinion of the reviewer, it is extremely unfortunate that the book emphasizes formal calisthenics as such an important part of the physical-training program. All exercise is classified into corrective, educational, hygienic, and recreative activities and these various phases are discussed from the standpoint of teaching as though the first three aims could best be secured through formal types of activities. The trend in physical education today is away from the old system of arm, leg, and trunk exercises and toward the

use of instinctive play activities. In my opinion it is very doubtful if the corrective, educational, and hygienic values are best secured through the type of formal exercises outlined in this book. The methods presented may be necessary under the crowded and congested conditions of city life but these conditions cannot be accepted as justification for such a system. Adequate playgrounds, play-spaces, playtime, in school and out, must be provided if physical education is to make any real contribution in education.

The discussion of the various forms of ptosis in relation to good posture and to their correction, is clear and definite. The two appendices give the general directions in classroom and gymnasium activities and a hygiene syllabus used in the public schools of New York during the time when the author was the director of physical education.

J. H. NICHOLS

Department of Physical Education, Ohio State University

DRUMMOND, MARGARET. *The Psychology and Teaching of Number.* Yonkers, New York, World Book Company, 1922. 126 pp.

This very readable little book was written and first published in England. It deals entirely with number as related to the pre-school portion of the child's life. Besides an introductory chapter we have chapters entitled "Psychological Stages," "Research Work," "Number and Intelligence," "Notation," "Games," "Mechanization of the Fundamental Processes," and "Conclusion."

In the introduction the author sets forth her point of view in a manner which will please even the most ardent proponent of children's rights. She says:

> Moreover, one may reasonably hope that we are progressing toward a time when children will no longer be the slaves of a bell, a time when they will no longer be squeezed or stretched to fit the same mould, a time when it will be realized that even the most biddable child cannot grow to order.

The author believes in "stages of growth," and enumerates six of them "through which every child passes, even although the rate of advance varies greatly." To attempt to teach a child upon a level which he has not yet attained is "perilous." Such teaching will only "confuse and retard" and will create "number imbeciles." Children must "grow before they are taught." "All our trouble arises from the fact that we *will* attempt to teach them while they are still at the animal level and before they have entered upon their heritage." Such teaching means "the gratification of our own sense of self-importance at the risk of forming in the child an 'inferiority complex' which will abide with him all his days, which will weaken his powers, and bring his finest endeavors to naught." This book contains much that suggests the "Culture Epoch Theory." This may offend the modernist, but it is so interestingly put that it will disturb no one who studied educational psychology previous to 1910.

The basis for the book is the careful study of a few individual children through the periods of their pre-school life. This is what "Research Work" means in Chapter III.

The author explains the frequent lack of correlation between a child's number work and his intelligence on emotional grounds. Children of special ability are "usually sensitive, and if they are hurried or puzzled or allowed to feel that they have made fools of themselves . . . . they will give up number and find themselves in some other sphere."

Such is the general spirit of the book, and the details are worked out accordingly. The book makes easy and interesting reading for anyone and should prove particularly valuable to the inexperienced primary or kindergarten teacher. One must always remember that the book was written with the English Infant Schools in view. Without this in mind, the inexperienced reader will misunderstand passages like the following quoted evidently with the approval of the author.

My own opinion is that if the founders of the policy which has directed elementary-school arithmetic until recently had sat down and deliberately planned a method of stifling the initiative and thinking power of the pupils, they could not possibly have invented one more efficient for the purpose than the one we have actually had.

Here the author is speaking of English schools. Let us hope that ours are not so bad.

WILLIAM J. OSBURN

State Department of Public Instruction,
Madison, Wisconsin

FREEMAN, FRANK N., AND DOUGHERTY, MARY. *How to Teach Handwriting.* Boston, Houghton Mifflin Company, 1923. 305 pp.

This manual enables the teacher of handwriting to get the pedagogical equipment she needs and furnishes the detailed exercises for her use together with the explanations which in the opinion of the authors are necessary to the intelligent use of these exercises.

The book consists of two parts. In the first part of the book the chief pedagogical and psychological principles which govern the learning process in writing are presented. Since they are largely based upon experiments and studies made in the laboratory of the University of Chicago, they furnish much thought-provoking information for the teacher of handwriting. This feature of the book should prove interesting and illuminating to teachers who are prone to follow traditional paths in the teaching of handwriting. Standard tests and penmanship scales and their uses are also fully discussed.

In the second part of the book the general principles laid down in the first part are applied in a series of exercises for the first six grades. The authors point especially to the following distinctive features: The exercises are arranged in the form of daily lessons; they are organized into a definite sequence; and a special set of aims for each grade is provided.

Four prominent features of the course are: (1) the omission of the general movement exercises in common use by teachers of handwriting (the oval and straight-line exercises); (2) the use of widely spaced exercises to develop the side-motion of the hand; (3) specific counting to stimulate and

regulate rhythmic practice of individual letters; and (4) the provision of correlated material adapted to the grade.

The first of these features is justified by the authors on the ground that "arm movement (muscular movement) is not recognized as being an important aim in the teaching of handwriting." This is something for teachers of handwriting to think about. Is the development of arm movement recognized as being an unimportant aim in the teaching of handwriting?

While the second and third features mentioned have been in more or less general use by professional teachers of penmanship for many years, the authors have so carefully woven them into the course as to provide an interesting method of procedure for the grade teacher.

This book is a valuable contribution to the cause of good penmanship because it contains a rich fund of ideas based upon scientific investigation and experiment, and a course in which these principles are applied.

C. C. LISTER

Brooklyn Training
School for Teachers

PINTNER, RUDOLPH. *Intelligence Testing.* New York, Henry Holt and Company, 1923. 406 pp.

This book presents a well-balanced statement of facts, principles, and present-day assumptions in intelligence testing; a quite complete historical sketch with due mention of the men who have contributed most; descriptions and evaluations of current methods and applications; and inspiring but not over-enthusiastic suggestions as to possible future developments.

The author announces in the Preface that the book is "an attempt to give a simple account of intelligence testing and the results that have so far been achieved by the testing movement." Considering the large amount of information presented, he has succeeded well in giving a *simple* account. He has, however, been forced to condense the material, until, in some parts, it is heavy reading for one not familiar with the field.

The Preface states that "the book is not a treatise on measurement in education or psychology," and that it "does not deal with educational tests." Had he likewise added that connections with educational psychology were to be made through references to the authorities in that field, he could have relieved the congestion of material by omitting occasional statements and the entire chapter on "The Inheritance of Intelligence." The author has, for the most part, kept within his field.

The general plan and arrangement of the book is logical, clear, and practical. The first fifty pages cover the historical aspects with well-placed emphasis on the psychological, educational, and sociological bases. The next forty pages are devoted to a thorough development of the concept of general intelligence. Pages 91 to 170 describe and illustrate the tests, both individual and group, which have been found to be of the most value. This section affords adequate concrete illustrations for the general and theoretical parts preceding, and its outline is such as to help to clarify definitions and perspective.

The final section, under the title of "The Results," and occupying more

than half of the volume, describes "the various fields in which mental tests have been successfully used," and sums up "the chief results in each field." The author's task here was not unlike that of Whipple fifteen years ago in the preparation of his *Manual of Mental and Physical Tests*, the material for which was scattered through all manner of publications or lying in crude manuscripts without having been evaluated or even sorted. Pintner's chapters and their bibliographies are a contribution that will be gladly welcomed. The only adverse criticism that the reviewer would offer here rests on the fact that in the chapters on "The Negro" and on "The Inheritance of Intelligence" he fails to find tests applied as actual working tools, but sees rather their use in profitable fields of research. These do not seem to fit well among the other chapters.

In writing on intelligence testing, one encounters at its worst the difficult question as to how much of theory should be presented as *fact* and how much as *open to dispute*. The non-technical reader is confused and discouraged if confronted continuously with controversial material and doubtfully presented data, but at the same time he must not be misled by dogmatic statements. The needs of those readers who are going deeply and thoroughly into the subject must also be kept in mind. Pintner is to be commended for his handling of disputed and uncertain features in a concise, clear, and interesting manner and without distortion, omission, or exaggeration.

R. H. SYLVESTER

Drake University

STENQUIST, JOHN L. *Measurement of Mechanical Ability*. New York, Teachers College, Columbia University, 1923. 101 pp. (Teachers College Contributions to Education, No. 130).

This is a pioneer investigation of general mechanical ability, which is defined by the author as "general aptitude in the management and manipulation of things mechanical." It may be distinguished from specific trade skill, since it involves the ability to deal with the things of everyday life rather than with the tools and processes of any one trade. Due to the increasing use of machinery this ability is of increasing value for its own sake, even if it were highly related to intelligence. The author points out that, in repeated experiments covering five years of testing in the public schools and in the army, the relation of general mechanical ability to intelligence when determined by performance tests is very low but positive. Correlations above 0.4 are exceptional. Consequently, under such a relationship, a good proportion of the group which now is destined to ultimate failure in academic school work may make good scores or even be above average in general mechanical ability.

General mechanical ability is shown to correlate highly with ability to do manual training shopwork, and so presumably to learn a trade. The use of the manipulative assembly mechanical tests will enable us to salvage the blind-alley-job group of pupils by pointing out to them their potentialities for the skilled trades. The tests may also be of much use in determining entrance qualifications to vocational schools—where evidently intelligence tests are all

but useless—and in advising boys in regard to prevocational courses. Presumably it ought to be worth while to attempt to teach mechanical principles in a prevocational class to boys who score low on this test in an attempt to keep them from growing up one-sided—"all intellectual" for instance. The reviewer recalls one eleven-year-old boy in the VIII-B or graduating class, with an I.Q. of about 150, who worked thirty minutes on the Stenquist Assembly Test and made zero, whereas the average of his classmates was 15 points.

The task, in the mechanical assembly performance tests, involves arranging without models ten disassembled mechanisms. The picture tests measure an element of general mechanical ability, but correlate fairly well with intelligence and hence do not have the "salvaging" value of the manipulative tests. Three forms of assembly test and two forms of picture test are now available. Full directions for administering and scoring the tests, age norms, and conversion tables are given in the volume. The tests are prepared for boys only.

This discovery of an aptitude which can be measured with high reliability by giving sufficient time for the testing, and which is but little related to the abstract academic type of intelligence, is of the utmost value for vocational guidance and education. The author might have pointed out, for instance, that the present trend towards requiring elementary-school graduation as an entrance prerequisite to vocational schools is a highly fallacious policy.

HERBERT A. TOOPS

Ohio State University

KOHS, S. C. *Intelligence Measurement.* New York, Macmillan Company, 1923. 312 pp.

The need for new tests or scales for individual testing is great, because there are very few which are valuable supplements to the Stanford-Binet. This is particularly true of performance and non-verbal tests. We may welcome this book, therefore, since in it the author discusses his Block Design Test, a well-constructed performance test. It has been carefully standardized and the results of this standardization are fully presented. The results are recorded by both the mental ages (the Binet) and the chronological ages of his subjects; the author seems to favor the former method. The two methods would seem to lead to about the same results, although we note that in the chronological-age standardization, all cases not within one year of normal grade for chronological age are discarded; this leads to an elimination of 33 percent of the cases. In the upper ages the selection of brighter children is obvious, and we therefore find a steady increase in the scores on the test from age five to age fifteen. Such a steady increase after the age of twelve is very rare in most individual performance tests. After age fifteen the increase is very irregular. Equivalent mental ages for scores are given up to age 19 years 11 months. Of course, such extension is necessary for calculating intelligence quotients, but the reliability of the mental ages and intelligence quotients which lie beyond the range of the standardization is very doubtful.

Kohs subjects his Block Design Test to all kinds of comparisons. He gives correlations with the Binet, the Trabue, and the Army Alpha test. The correlation between the Binet mental age and raw score on the test is 82, and

between the Binet mental age and weighted score 83. One is, therefore, not impressed by the necessity for weighting and particularly not according to the ratio 4 : 2 : 1, simply because the correlations with chronological age for scores, time, and movements were 71, 38, and 20, respectively.

The test should be of particular value to the clinical psychologist who feels the need of additional non-language tests. As the author points out, both language and performance tests are needed, and the latter often give us much insight into what the author calls the "loquacious" and "silent" types of children. Furthermore, watching a child at work on a performance test will often give the examiner an insight into certain aspects of his character which cannot be obtained during the question and answer examination of the Binet.

This, then, is the gist of the book, and the reviewer has nothing but commendation for the author's work in giving us a new and well-standardized performance scale. Unfortunately, in the opinion of the reviewer, the author has appended before and after and interpolated between the straight story of his experimental work the discussion of a host of other related topics. Thus the book begins with an elaborate philosophical, one might almost say metaphysical, discussion of analysis and synthesis. There are far too many quotations here. We are left guessing as to the reason for this elaborate chapter until we come, much later in the book, to the author's definition of intelligence as "the ability of an individual to analyze and synthesize." Does it need all this massive superstructure to vindicate or explain this definition of intelligence? Again, the general story of the work is halted for an interesting discussion of standardization technique, which might well have formed a separate article in a technical journal. There is also a discussion of the average mental age of adults and the problem of mental growth, not to mention four appendices.

All of these topics are important, and the author has interesting contributions to make to each of them. What irritates the reviewer, and probably will irritate the average reader, is the way in which they have been put together to form a book which bears the comprehensive title *Intelligence Measurement*. A better title for the book would probably have been its present sub-title, "A Psychological and Statistical Study Based upon the Block-Design Tests."

R. PINTNER

Teachers College, Columbia University

THORNDIKE, EDWARD L. and others. *The Psychology of Algebra.* New York, The Macmillan Company, 1923. 483 pp.

This book presents a number of studies carried out by a number of people, headed by E. L. Thorndike, in the Institute of Educational Research of Teachers College, Columbia University. Credit is given the Commonwealth Fund for a grant making possible the investigations that are here reported.

The eighteen chapters are concerned with various topics relating to the teaching of algebra, commencing with studies of the select group of boys and girls who enter American high schools, and proceeding through detailed

studies of the nature and constitution of algebraic abilities to treatment of such details as the equation, problem solving, drill, and measurement of algebraic abilities.

Throughout the book, the point of view of Thorndike's psychology of learning is dominant. The dynamic aspect of mind as a system of connections between stimulus and response is continually emphasized. Learning is treated as the formation of such connections or elementary habits; thought and reasoning are regarded as effective organization of these habits.

The classroom teacher of high-school mathematics is certain to profit by reading the discussion of the abilities most worth while acquiring in a year's course in algebra, regardless of his complete agreement with the topics listed. The detailed suggestions concerning particular items of learning to be added or omitted, and the discussion of the learning of meanings and procedures in algebra are sure to be significant in helping to bring about desirable changes in classroom teaching.

Frequent reference is made to the work of the National Committee on Mathematical Requirements; and certain chapters, such as the ones on "The Uses of Algebra" and "The Interest of Pupils in Algebra," supplement portions of the Report of the National Committee.

Instruction in mathematics in our high schools today probably engages more careful thinking concerning its problems than does any other subject in the program of studies. The present book makes contributions that will cause it to be welcomed by all teachers of mathematics.

J. A. Foberg

State Department of Public Instruction
Harrisburg, Pennsylvania

Wood, Ben D. *Measurement in Higher Education.* Yonkers-On-Hudson, New York, World Book Company, 1923. 337 pp.

This book gives us a stimulating discussion of the measurement problem as it presents itself to our institutions of higher learning. The results of the use of mental and educational tests with students at Columbia University during the past few years supply the experimental data upon which the discussion is based. While the book deals with a specific experiment, the treatment and conclusions are of general value. They will prove helpful to instructors and administrators both in secondary schools and colleges. The writer shows that the same measurement methods which have proved their value in lower schools are likewise of practical use in higher education.

The first part of the book deals with the problem of determining the value of the Thorndike Intelligence Examination for High-School Graduates as a criterion for admission to Columbia College. This question is answered by finding the correlations between the two-year scholarship scores of students in Columbia College and their scores in three different criteria for admission, as follows:

Thorndike Intelligence Scores	.672
New York State Regents' Examination Marks	.644
Secondary School Marks	.262

If scholarship is the best measure of college success, the intelligence score is a better criterion for admission to college than either of the other two.

Furthermore, the data show that the correlations of scholarship scores for different periods of time in no case exceed .724. The intelligence examination, then, predicts college success very nearly as well as college success, as expressed by the best available index, predicts itself from one semester or year to another. Accordingly, any hope of improving the intelligence test's predictive value depends upon improvement in the reliability of college grades.

A large part of this unreliability is due to the indefinite and variable bases of grading and to poor technique of administering examinations. The latter half of the book attempts to show how this condition is remediable. After presenting those principles of educational measurement which should be observed in the construction of tests and examinations, the writer devotes the remaining chapters of the book to a discussion and illustration of this new type of examination in various subjects and courses common to colleges and secondary schools.

C. E. Erffmeyer

North-Western College

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Doctor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

The World Book Company has recently announced the Blackstone Stenographic Proficiency Test in five forms. Norms are given for varying amounts of instruction at five-month intervals covering a period of thirty months.

Data from Cornell University show that 30 percent of the farmers who earn an annual income of over $1,000 are high-school graduates, while but 5 percent of those earning over $1,000 had a common-school education or less.

A recent mimeographed bulletin from the Department of Research of Duluth shows that for the five-year period, 1919-20 to 1923-24, the largest increase in cost of instructional service in any department has been 54.9 percent, and the least, 34.9. The increase for the entire school system is 47.5 percent.

The second number of the *Maryland Schools Bulletin* for the present year is a fit companion for Number I. This number (October) is entitled, "Does the Helping Teacher Really Help the Rural Schools?" It presents with a large amount of suggestive detail a year's work in rural supervision.

The National Honesty Bureau which is maintained by the National Surety Company has issued the *Honesty Book* which might be termed a teacher's handbook in methods of character education, especially in this phase.

Those interested in vocational guidance, if they have not already done so, should make the acquaintance of the *National Vocational Guidance Association Bulletin*, published by the Bureau of Vocational Guidance, Graduate School of Education, Harvard University. The bulletin brings together concrete material from the experience of those at work in this field.

School and Society for November 24 notes the activity of the Indianapolis Chamber of Commerce in providing scholarships and other forms of aid for the graduates of superior ability whose parents cannot afford to send them to

college. It is also organizing a system of educational guidance for use in the city, and is publishing pamphlets on occupations based on scientific investigations.

The annual report of the chief medical officer of the Board of Education in England in 1922 shows that 46 percent of the children in average daily attendance were given medical examinations. Ninety-three percent of the education authorities provided school clinics; and 92 percent made provision for supplying spectacles for children having defective sight. About 42 percent of those examined were defective in some degree.

Superintendent E. B. Sellew of Middletown, Connecticut, reports that 1.5 percent of the children in the elementary grades are young-rapid; 33 percent, normal in age and progress; and 26 percent, over-age and retarded. In the high school 15.4 percent are young-rapid; 22.6 percent are normal in age and progress; and 12 percent, over-age and retarded.

We note with interest in the *High School Journal* of November that Doctor M. R. Trabue is standardizing a test in the appreciation of English prose. The test is devised along the same line as the Abbott-Trabue exercises in judging English poetry. The test consists of 30 prose paragraphs, each printed in three versions, and ranked as best, average, and poorest.

Mr. W. L. Brown is in charge of a regularly organized Department of Reference and Research in the New Trier Township High School, Kenilworth, Illinois. This is the first information we have received of an organized department of research in a high school. If there are others, we shall be glad to learn of them either from the director or from any one else who may know of such an organization.

Superintendent W. H. Holmes, Mt. Vernon, New York, in his annual report, *The Individual Child in the Schools*, recounts the work which is being done in that school system in individualizing instruction. Those who are interested in the field of individual instruction will find here definite information and the opinions expressed by teachers and principals under whom the work is being carried on.

The United Parents' Associations of New York City recently launched a campaign for $50,000. This money is to be used to conduct the business of the organization, to employ experts on various educational problems, to promote education, and to conduct such other activities as come within the association's scope. The purpose of the United Parents' Association is to combine in one inclusive organization the strength of the individual associations. It is absolutely non-political.

On November 10 the schoolmen of Western Pennsylvania met at the University of Pittsburg and formed an association of school administrators.

Superintendent Ben Graham of New Castle was elected president, and C. W. Stone, Munhall, was elected secretary-treasurer. The organization expects to meet two or three times each year. The first program was devoted to community building programs with Doctor Engelhardt of Columbia University as the principal speaker.

The recent special session of the Arkansas Legislature passed a law requiring the teaching of the Constitution of the United States in all high schools, academies, colleges, and universities throughout the State, and making the completion of such a course a pre-requisite for graduation from any of these institutions. All applicants for teaching are now required to pass a certain examination on this subject before being licensed to teach in the State.

A bulletin on the *Size of Classes in the Elementary Schools of Duluth* shows that the median for regular classes in grades one to eight varies from thirty-three to thirty-seven. The greatest range of individual classes, however, is from one class with an enrollment of forty-eight in the sixth grade to one with an enrollment of eighteen in the eighth. These data are for November 1 of the current year.

A comparison of the size of classes from 1917 to 1923, inclusive, shows that Duluth has been able to keep the median class almost constant, while the number of very large classes has been gradually reduced.

Superintendent J. R. Patterson of Bucyrus (Ohio) has been furnishing us for some time with regular contributions in the way of survey bulletins. Apparently the work is growing since Volume I, Number 5, October, 1923, has just appeared in four parts. Part I is concerned with age-grade and brightness data; Part II, reading, spelling, and handwriting; Part III, arithmetic and geography; and Part IV, English composition. The bulletins, in addition to presenting survey results, also contain suggestions to teachers and supervisors for increasing the effectiveness of teaching. Part IV on English composition is especially full of helpful suggestions.

The World Book Company announces the publication of the Schutte Scale for Rating Teachers. The author, T. H. Schutte, director of the training department, Northern Arizona Normal School, states that the scale is a device to guide and assist one in the rating of teachers in those qualities which make for success in teaching. The five general heads under which a sum total of 86 questions are asked are as follows: Personal and Social Qualities; Cooperative Qualities; Leadership; Scientific and Professional Attitude; and Teaching Ability.

In a recent address before the New York Society for the Experimental Study of Education, Dr. Ira S. Wile, former member of the Board of Education, stressed the importance of physical examination of the children of pre-school age. He drew attention to the fact that physical defects have

been revealed in large numbers in children under school age, and that a majority of these defects can quite easily be remedied. He convincingly argued that this work should be done during the half-year previous to the child's school entrance. Such remedial work would not only be of direct material assistance to the children, but would aid the school in proper classification and care after entrance.

A very interesting mimeographed bulletin, showing school costs for 1921-22 for each county of North Carolina, has been received from the State Department. Costs are grouped under four major heads: Administration, Operation and Maintenance, Teaching and Supervision, and Outlay Payments. For each county is given the per pupil expenditure under each of these heads, and the total expense per pupil. For the state as a whole, 88.6 percent of the current expense is spent for teaching and supervision; 8.4 percent on maintenance and operation; and 3 percent on administration. When the outlay payments are included, the percent for teaching and supervision is 63.6; for operation and maintenance, 6; and for administration, 2.1.

The committee on the American Peace Award, created by Edward W. Bok and offering $100,000 for the "best practicable plan by which the United States may cooperate with other nations looking toward the prevention of war," will undertake a nation-wide referendum on the plans selected by the Jury. A large number of newspapers and magazines are carrying the text of the winning plan and a digest for a ballot which will enable the reader to forward to the Committee his approval or disapproval of the plan. Several hundred organizations and institutions are sending the winning plan with ballot to each of their members for a vote. Any number of copies of the winning plan accompanied by ballots may be secured without charge by writing to the American Peace Award, 342 Madison Avenue, New York City.

The following facts relative to vacation courses for teachers are gleaned from a report in the October number of the *American School Master*. The data were gathered by the Flint (Michigan) Teachers Club for the school year 1921-22. Of 31 cities of a population of 75,000 to 125,000 throughout the United States, 16 made no allowance for a course pursued in summer school; 27 gave no compensation for attending national meetings or travel; 17 granted no allowance for correspondence or extension courses; and 27 made no allowance for social service work. Of 19 Michigan cities replying to the same questionnaire, 11 gave no allowance for attending summer school; 13, no allowance for travel; 11, no allowance for correspondence or extension courses; and 17, no allowance for social service work.

The difference between the number given and the number of cities reporting need not be interpreted as a positive answer since in a number of cases it was a qualified affirmative or negative which heavily weighted the results.

During the election campaign in England last month the Independent Labor Party issued a statement on "How a Socialist government would deal

with education.'' Certain excerpts from the statement may be of interest to our readers.

1. The true aim of education should not be to mould opinion, to turn out good workmen, good soldiers, or even good Socialists. Education is not propaganda, it is the development of the mind.

2. A Socialist government would consider the provision of the fullest education for every child to be its supreme task. It would not consider expenditure upon education to be wasteful, but would look upon it as a highly profitable investment.

3. The present educational system is gravely inadequate and unsuited to its purposes; classes are too large; physical needs of the children are neglected; children leave school too early; teachers have little freedom and are badly paid; and the curriculum is cramped.

4. A Socialist government would raise the school-leaving age to 18, and a university course would be available to all young persons capable of profiting from it. A Socialist government would not hesitate to economize in war and expend on education.

The teachers' bulletin of Highland Park (Michigan) for November 5 reports a study of high-school failures during the first semester of the school year 1922-23. There were approximately 450 pupils who had failed in one or more subjects, the total number of failures being 527 for boys and 343 for girls. Each teacher in the high school was asked to check up on the causes of failure, and for the 870 failures 1,802 reasons were given. These were distributed as follows:

Cause of Failure	Percent
Laziness (so-called)	20
Dullness	16
Poor preparation	16
Unwillingness to cooperate	11
Absences excused	11
Absences unexcused	7
Outside activities	3
Wrong classification	2
Too easy schedule	1
Miscellaneous reasons	13

The teachers also gave various remedies for the failures mentioned.

Remedies	Number For Whom Suggested
Closely supervised study	40
More time	28
Attempt fewer studies	6
Discipline at home	4
Repeat work	3
Have better ancestry	3
Complete change of environment	2
Less home work	1

It is noteworthy that the teachers' suggestions of remedies apply to only about twenty percent of the total number of failing children involved. This presents the matter from the teachers' standpoint alone. We wonder what replies the failing pupils would have given had they been asked the same question. Would poor teaching, unsympathetic attitude of teachers, and too long

assignments—reasons which place the responsibility upon the school rather than upon the child—have found a place among the reasons given?

Report of the Progress of the Reading Committee

The Committee on Reading, appointed by Commissioner Tigert in January, 1923, recently reported substantial progress in its work. It will be recalled that the committee was asked to make a study of the debatable issues in the field of reading instruction and to make recommendations based on experimental evidence and expert opinion for the aid of teachers and supervisors. The first meeting of the committee was held in Cleveland during the latter part of February in connection with the meeting of the Department of Superintendence. At that time a general plan of procedure was adopted. The first step taken was to secure a list of the problems of reading instruction which are causing most difficulty among teachers and supervisors. In this connection, reports were secured from fifty-seven cities. With this information at hand, the committee met in Chicago in May for a deliberate study of the problems which had been suggested. During the course of the conference, it became evident that most of the problems could be organized about seven or eight issues of major importance. Accordingly, the plan was adopted of assigning a large problem to each committee member who was authorized to select a group of people to aid him in the investigation of the problem and the preparation of a report. The names of the chairmen of these committees and the problem assigned to each follow:

1. Types of Reading of Large Social Value—Professor S. A. Leonard, University of Wisconsin.
2. An Outline of a Program of Instruction in Reading—Dean W. S. Gray, University of Chicago.
3. Special Types of Reading Activities in Content Subjects—Miss Estaline Wilson, Assistant Superintendent of Schools, Toledo, Ohio.
4. Appropriate Materials of Instruction—Professor Ernest Horn, The University of Iowa.
5. How to Develop Independence in the Recognition of Words—Miss Frances Jenkins, The University of Cincinnati.
6. Individual Differences, Tests, and Remedial Treatment—Miss Laura Zirbes, The Lincoln School, New York City.
7. How to Put Across a Progressive Program of Reading Instruction—Superintendent F. W. Ballou, Washington, D. C.

During the summer and autumn of 1923, the committees carried forward their work although numerous difficulties were encountered owing to the illness of various members. At present, however, practically every committee has prepared a preliminary report. In order to complete the work assigned to the Reading Committee, it will be necessary to have two general conferences: the first, in order to reach an agreement with regard to the fundamental issues which affect all sections of the final report; and the second, to check the validity of each item and recommendation which is included. At the present time, it is planned to postpone the final conference until after a summary has been made of all the scientific material on reading. Provision has been made through a subsidy from the Commonwealth Fund for such a summary which will be prepared by Dean William S. Gray of the University of Chicago during April and May, 1924.

Educational Research Association

(HARRY A. GREENE, *Secretary and Editor*)

The *Educational Research Association* will hold two open meetings, one on Tuesday afternoon, Feb. 26, at 2:00 o'clock in the Tiger Room of the Sherman Hotel and the other on Wednesday morning, Feb. 27, at 9:15 o'clock in the Banquet Room of the Auditorium Hotel. The Wednesday morning meeting is a joint meeting of the Educational Research Association and the Society of College Teachers. The topic for discussion at the Tuesday afternoon meeting is Improvement of Instruction in Arithmetic. Mr. P. R. Stevenson of Ohio State University will speak on A Reading Test in Arithmetic; Director C. H. Judd, of the School of Education, University of Chicago on Laboratory Analysis of Arithmetic; Dr. W. J. Osburn of the State Department of Public Instruction, Madison, Wisconsin, on Amount of Drill Revealed Through the Analysis of a Textbook in Arithmetic; Professor F. B. Knight of the State University of Iowa on Transfer in the Narrow Mental Function of Adding Fractions; Dr. B. R. Buckingham of the Ohio State University on Individualizing Instruction on the Basis of Tests; and A. B. Wight of the Chicago Public Schools on Improvement in the Fundamentals of Arithmetic Through Tests.

At the meeting Wednesday morning, W. W. Charters will speak on Technique of Trait Analysis; W. H. Hughes, Director of Research and Guidance, Pasadena City Schools, California, on Organized Personnel Research and Its Effect on High School Practices; Dean W. S. Gray of the School of Education, University of Chicago on Case Studies of Reading Deficiencies in Junior High Schools; J. Freeman Guy, Director of Research and Measurement, Pittsburgh, Pennsylvania, on Diagnosing the Individual Case; Dean S. A. Courtis of Teachers College, Detroit, Michigan, on Relation of Rate and Quality in Educational Measurements; J. K. Norton, Director of the Research Division of the N. E. A. on A General Survey of the Curriculum Situation; and Professor F. N. Freeman of the University of Chicago on What Can Motion Pictures Contribute to Education?

The executive meeting for the members of the association will be held Thursday morning, February 28 at 9:15 o'clock in the Crystal Room of the Sherman Hotel and the banquet of the association will be held at 6:15 Thursday evening, February 28, in the Crystal Room of the Sherman Hotel.

The program of the executive session follows:

Thursday Morning, February 28, 9:15 o'clock
Crystal Room, Sherman Hotel

STATISTICAL AND EXPERIMENTAL STUDIES

Method of Coordinating Results of Mental Tests and Achievement Scales in Surveys of School Achievement and Progress (Twelve Minutes)
M. J. Van Wagenen, College of Education, University of Minnesota, Minneapolis, Minnesota

Discussion Led By (Five Minutes)
Joseph Peterson, George Peabody College for Teachers, Nashville, Tennessee

A Graphic Device for Representing Test Scores Over a Period of Years (Ten Minutes)
M. R. Trabue, School of Education, University of North Carolina, Chapel Hill, North Carolina
Discussion Led By (Five Minutes)
W. W. Theisen, Assistant Superintendent of Schools, Milwaukee, Wisconsin
A Graphic Method for Obtaining the Coefficient of Correlation (Twelve Minutes)
E. R. Wood, Kansas State Teachers College, Emporia, Kansas
Discussion Led By (Five Minutes)
Raymond Franzen, University of California, Berkeley, California
Skewness of Distribution of Reading Scores and the Supervision of Instruction (Twelve Minutes)
H. S. Will, Statisticiaa, Toledo, Ohio
Discussion Led By (Five Minutes)
Bessie Lee Gambril, School of Education, Yale University, New Haven, Connecticut
The Criteria for Planning and Judging Spelling Tests (Twelve Minutes)
Ernest Horne, School of Education, State University of Iowa, Iowa City, Iowa
Discussion Led By (Five Minutes)
R. L. Morton, Ohio University, Athens, Ohio
An Experiment on the Merits of the Individual Plan of Instruction (Twelve Minutes)
C. W. Washburne, Supt. of Schools, Winnetka, Illinois
Discussion Led By (Five Minutes)
H. R. Rugg, Lincoln School, Teachers College, New York